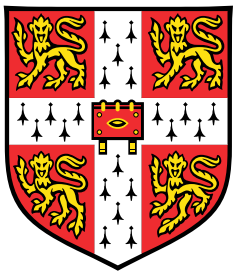
City-scale eco-routing and pavement eco-maintenance scheduling for CO_2 mitigation



Bingyu Zhao

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Churchill College June 2019

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Bingyu Zhao
June 2019

Acknowledgements

I would like to thank Professor Kenichi Soga for being the greatest supervisor. Since when I nervously had the first Skype interview with him five years ago, Kenichi has always been extremely kind, patient and supportive. I enjoy discussing about my research with him and can always learn from his attitude when facing difficulties in my work. Kenichi encourages me to think outside of my comfort zone and it would be impossible to have this thesis without him.

Sincere thanks to Dr Elisabete Silva for being my advisor. Elisabete is always ready to listen about my research and offers me great technical guidance. She also cheers me up when I feel confused. Also to Dr Ruchi Choudhary, who kindly took me in as a student when Kenichi moved to Berkeley. The idea of one of the key parts of this thesis (the spatial pavement degradation model) originates from Ruchi.

I would also like to thank my examiners Professor Peter Guthrie and Dr Nuno Pinto for spending time in reading my thesis and providing me with helpful and encouraging feedback. I hope to carry on this work incorporating their advice, making the viva as the start of the next stage of research.

To Gerry and Krishna, who are the first people I want to share with when I face setbacks or make progress. I am so happy to have them as collaborators and I hope they would enjoy it as well.

And to my other academic brothers and sisters: Vanessa for being my neighbour and best friend who I can talk about everything and keep each other motivated. Tsubasa, Hansini, Yuze, Ray and May, Yi and He, Zili and Lulu, Linqing and others for their friendship.

I am grateful to have the best colleagues at CSIC, the Turing and UC Berkeley for expanding my horizons with knowledge and support.

And my friends Mingfei and Ben, Yuanjun, Tiantian, Miki, Paula, Alisha, Gabi and Elsa for all the good times and nice foods.

Finally, I must thank my parents for their love and understanding. I feel sorry sometimes for being so far away. You are my best friends in life and I feel so lucky to have parents like you!

Abstract

Responsible for more than one-sixth of the world's CO_2 emissions and growing, the road transportation is an unavoidable component in tackling the global carbon reduction challenge. Within the subdivisions of infrastructure management and traffic management, different carbon mitigation approaches have been discussed to contribute to a greener transportation system. However, given the complex interactions that exist between the road users and the infrastructures, certain carbon mitigation proposals are best evaluated within a comprehensive environment to ensure the correct consideration of the interactions, as well as the comparability of the results.

It is the aim of this thesis to develop such a unified framework to reflect interactions and outcomes of different CO_2 mitigation approaches from both the road infrastructure and the traffic mobility sides. This allows combining and comparing the effectiveness of a single or multiple carbon mitigation approaches across different perspectives. Specifically, on the traffic operation side, a mesoscopic traffic model is adopted for simulating drivers' route choices with varying percentages of travellers to choose the eco-friendly routes. From the infrastructure asset management perspective, a city-scale pavement degradation model is built and utilised in testing pavement maintenance scenarios. San Francisco is chosen as the case study area due to the availability of various traffic mobility and infrastructure condition data.

In Chapter 3 of the thesis, the traffic simulation module is developed that implements the efficient macroscopic road link-level speed-flow relationship while retaining detailed origin, destination and departure hour information for each individual trip. The traffic simulation model uses a highly detailed network representation for the study area and has the hourly traffic demand informed by Traffic Network Companies (TNC) data. The model is capable of capturing the spatio-temporal variations in traffic distributions. In addition, in Chapter 4, assumptions are also tested as for the extent that the availability of real-time traffic information affects the travellers' behaviours and model results.

As ageing pavements induce additional carbon emissions, in Chapter 5, a city-scale pavement degradation model is proposed based on 20 years of survey data. After comparing three model forms (non-spatial categorical, non-spatial individual road based and spatial

hierarchical models) and two independent predictors (pavement age, cumulative traffic load), the spatial model with age as the predictor is found to give the best overall performance in terms of model fitting and complexity. As a result, it is used later in this thesis for degradation forecasting and maintenance planning.

The traffic simulation and the pavement degradation models are joined together in Chapter 6 to test the carbon mitigation scenarios, including the eco-friendly route selection (eco-routing) and eco-friendly pavement maintenance scheduling (eco-maintenance). Interactions between the road users and the pavement management occur when: (1) pavement maintenance site selection is based on both pavement roughness and traffic volume (the eco-maintenance case). (2) The renewed pavement condition, with a smoother surface and reduced emission factor, becomes part of the route selection criteria of the drivers (the eco-routing case). It is found that the outcomes of eco-maintenance are sensitive to a variety of factors, including the budget level, the pavement degradation rate as well as the maintenance quality. The eco-routing approach tends to shift travellers to the local network but is effective in reducing the overall emissions. However, the reinforcing interactions between these two strategies are the most noticeable only when both eco-routing and eco-maintenance strategies are enforced to an extreme.

Through the simulation of city-scale traffic and infrastructure dynamics, it is able to quantitatively compare carbon mitigation scenarios and understand how an action from one specific perspective ripples through the transportation system. Also, sensitivity tests suggest limiting conditions for each approach to make a difference. This research highlights the need to include combined simulations in certain cases and such results are expected to give confidence to decision makers as for the potential induced demand or other secondary effects whose influences extend beyond a single sub-system.

Table of contents

List of figures

List of tables

Nomenclature

Roman Symbols

a_0, a_1, a_2 unitless regression coefficients in the CO_2 emission model of Wang et al. [241]

age_{ik} age of pavement segment i in year k

b_i unitless regression coefficients in the CO_2 emission model of Barth and Boriboonsomsin [15]

T relaxation parameter in the car following model

D measurement distance in speed calculation

$D_{i,2015}$ number of trips ending in District i in 2015

$D_{i,2050}$ number of trips ending in District i in 2050

$ESAL_{ik}$ Equivalent Single Axle load (ESAL) of pavement segment i in year k

FFS free flow speed in the road capacity calculation

f_p city-specific correction factor for the BPR volume delay curve

g gram(s)

$g_{i,D}$ annual growth rate for trips ending in District i

$g_{i,O}$ annual growth rate for trips originating from District i

g/C effective green ratio of the traffic light in the road capacity calculation

$\%HV$ heavy vehicle ratio in the road capacity calculation

k traffic density

S_0, S_n Pavement condition as a vector in Year 0, Year n.

T Markov Transition Probability Matrix

mph mile(s) per hour

N number of vehicles

n_i the numbers of neighbours that street i has in the degradation model

N_p number of all probes that go pass link l

N_{th} number of through lanes in the road capacity calculation

$N(x,t)$ vehicle counts at space (x) and time (t)

$O_{i,2015}$ number of trips originating from District i in 2015

$O_{i,2050}$ number of trips originating from District i in 2050

PCI_{ik} Pavement Condition Index (PCI) of pavement segment i in year k

p_D effective number of parameters in the calculation of DIC

q traffic flow rate

s saturation flow rate for the through movement in the road capacity calculation

sp abbreviation of the spatial model, including $sp-age$ and $sp-logESAL$

T elapsed time in flow calculation

t time, in the vehicle conservation equation

t_f free flow travel time

$t_{f,l}$ free flow travel time on link l

t_i time take for vehicle i to go through a measurement distance

t^A_{min} the free-flow travel time in explaining the queue-based model in MATSim

t_t travel time with traffic on the link

$t_{t,l}$ aggregated probe measured link-level travel time t for link l

$t_{t,l,p}$ individual probe measured link-level travel time t for link l measured by probe vehicle p

u_i or $\boldsymbol{u}$ spatially unstructured deviation from the mean of the age or traffic load effect β in the pavement degradation model, for street i or all streets

v speed; or average speed

$v_{follower}$ speed of the following vehicle in the car-following theory

v_i, v_j or $\boldsymbol{v}$ spatially structured deviation from the mean of the age or traffic load effect β in the pavement degradation model, for street i, j or all streets

v_{leader} speed of the leader vehicle in the car-following theory

v_s space mean speed

x postion, in the vehicle conservation equation

x_{ik} the explanatory variable (age or cumulative ESAL) corresponding to the k-th observed PCI at street i

$x_{t,l,p}$ travel delay per unit distance experienced by driver/probe p on link l

$\boldsymbol{y}$ observations in the Bayesian model

y_{CO_2} CO_2 emissions (grams per mile)

$y_{\mathrm{CO}_2,level}$ CO_2 emissions at a levelled ground (grams per mile)

$y_{\mathrm{CO}_2,slope}$ CO_2 emissions on a sloped road (grams per mile)

y_{ik} the k-th observed PCI at street i

c an anticipation term in second-order macroscopic flow models

T_{elp} relaxation time to adjust the speed in second-order macroscopic flow models

$V_e(k)$ speed at steady state provided by the fundamental diagram in first-order models

Greek Symbols

α BPR volume delay curve coefficient; or global average intercept shared by all streets in the degradation model

α_i intercept in the linear relationship between log ESAL to PCI

β BPR volume delay curve coefficient; or global average age or traffic load effect shared by all streets in the degradation model

β_i slope in the linear relationship between log ESAL to PCI

η_{ij} or η_{ik} the mean of the normal distribution used to obtain the j or k-th observed PCI at street i

γ variability term in the multiplicative error structure for $x_{t,l,p}$, the travel delay per unit distance experienced by driver/probe p on link l

$\mu_{t,l}$ true average link-level travel time t given by the BPR function for link l

$\mu_{xt,l}$ mean of $x_{t,l,p}$, the travel delay per unit distance experienced by driver/probe p on link l

π shape parameter used in the *Gamma* distributions of γ and $x_{t,l,p}$; probability function in explaining the Bayesian theorem

$\boldsymbol{\psi}$, ψ_h hyperparameters of the degradation model, $= \{\tau_g, \tau_\xi, \tau_u, \tau_v\}$

ψ_h marginal posterior of a hyperparameter in the pavement degradation model

$\sigma_{t,l}$ variability in individual vehicle's link-level travel time t for link l

σ_ξ standard deviation parameter of the truncated normal distribution

τ_g the precision (inverse of variance) of the normal distribution used to obtain the k-th observed PCI at street i

τ_u precision of u_i in the Besag model

τ_v precision of v_i in the Besag model

τ_ξ mean value of the street specific variations in intercept

$\boldsymbol{\theta}$, θ_w parameters of the degradation model, $= \{\alpha, \beta, \boldsymbol{\xi}, \boldsymbol{v}, \boldsymbol{u}\}$

θ_w marginal posterior of a parameter in the pavement degradation model

$\widetilde{\pi}_G$ Gaussian approximation of a probability distribution π

ξ_i or $\boldsymbol{\xi}$ street specific variations in intercept for street i or all streets

Acronyms / Abbreviations

A road functional class: arterial

AAD-CO_2 annual average daily CO_2 emissions

AADT annual average daily traffic

AAD-VHT annual average daily vehicle hours travelled

AAD-VKMT annual average daily vehicle kilometres travelled

AASHO American Association of State Highway Officials

AASHTO American Association of State Highway and Transportation Officials

ADOT Arizona Department of Transportation

AIMSUN a commercial microscopic traffic simulation programme

API application programming interface

ARZ Aw-Rascle-Zhang model, a second order macroscopic traffic propagation model

ATIS advanced traveller information system

Ave avenue

BB Bay Bridge

BEAM Behaviour, Energy, Autonomy and Mobility for simulating transportation and electric system

Blvd boulevard

BPR Bureau of Public Roads

CA arterial roads with asphalt concrete overlaid on Portland cement concrete

Caltrans California Department of Transportation

CAR conditional autoregression

CATS Chicago Area Transportation Study

C pavement surface type: asphalt concrete overlaid on Portland cement concrete; or Road functional class: arterial

CC collector roads with asphalt concrete overlaid on Portland cement concrete

CEN Comité Européen de Normalisation

CMEM Comprehensive Modal Emissions Model

COV coefficient of variation

CR residential roads with asphalt concrete overlaid on Portland cement concrete

CSD3 Cambridge Service for Data Driven Discovery

DAG directed acyclic graph

DEM digital elevation map

DIC deviance information criterion

DTA dynamic traffic assignment

DynaMIT Dynamic Network Assignment for Management of Information to Travelers. A mesoscopic traffic simulation software programme

DYNASMART DYnamic Network Assignment Simulation Model for Advanced Road Telematics. A mesoscopic traffic simulation software programme

EIA United States Energy Information Administration

ESAL equivalent single axle load

EU European Union

FCD floating car data

Fwy freeway

GGB Golden Gate Bridge

GHG greenhouse gas

GMRF Gaussian Markov random field

GPS global positioning system

GTFS general transit feed specification

GWR geographically weighted regression

HCM Highway Capacity Manual

HDM-4 Highway Development and Management 4

HIPS Highway Investment Programming System

HMEP Highway Maintenance Efficiency Programme

HOV high occupancy vehicle, a vehicle with a driver and one or more passengers

HPC High Performance Computing

Hwy highway

iCAR intrinsic conditional autoregression

INDY INterative DYnamic, a traffic assignment model

INLA integrated nested laplace approximation

INRIX a private company in the US providing location-based data and analytics

IQR inner quartile range

IRI International Roughness Index

ITS intelligent transportation systems

JSON JavaScript Object Notation

kip kilo pound, a US customary unit of force equalling 1000 pounds-force

kN kilo Newton, International System of Units derived unit of force

LSFCD large scale floating car data

LBNL Lawrence Berkeley National Laboratory

LCA life cycle analysis/assessment

LOS level of service, a traffic congestion measure, where A is the best and F is the worst condition.

LTPPM Long Term Pavement Performance Model

LWR Lighhill-Whitham-Richards model, a first order macroscopic traffic propagation model

MARPLE Model for Assignment and Regional PoLicy Evaluation

MATSIM or MATSIM-T Multi-Agent Transport Simulation Toolkit. An agent-based mesoscopic traffic simulation software programme

MCMC Markov Chain Monte Carlo

M-E mechanistic-empirical

MEPDG Mechanistic-Empirical Pavement Design Guide

METANET a macroscopic simulation program for motorway networks

MFD macroscopic fundamental diagram

MLE maximum likelihood estimation

MPD mean profile depth, a measure of pavement macroscopic texture

M&R maintenance and rehabilitation

MTC Metropolitan Transportation Commission in the San Francisco Bay Area

NCHRP National Cooperative Highway Research Program

OA arterial roads with asphalt concrete overlaid on asphalt concrete

OD origin-destination

O pavement surface type: asphalt concrete overlaid on asphalt concrete

OR residential roads with asphalt concrete overlaid on asphalt concrete

OSMnx Python tool for retrieving, modelling, analysing and visualising street networks and spatial data from OpenStreetMap

OSM OpenStreetMap

PA arterial roads with Portland cement concrete

PARIS Performance Analysis of Road InfraStructures model

PCI Pavement Condition Index

PC collector roads with Portland cement concrete

PEV plug-in electric vehicles

P pavement surface type: Portland cement concrete

PR residential roads with Portland cement concrete

$Q1$ 1st quartile

$Q3$ 3rd quartile

QL query language

RMSE root-mean-square errors

R road functional class: residential. Or the software R

SAR simultaneous autoregression

SFCTA San Francisco County Transportation Authority

SFMTA San Francisco Municipal Transportation Agency

SFO San Francisco International Airport

SF San Francisco. Case study area of this study

SR 1 California State Route 1

SS sub-step (used in figures such as Figure 3.14)

STA static traffic assignment

St street

SUMO Simulation for Urban MObility

TAZ traffic analysis zones

TIGER Topologically Integrated Geographic Encoding and Referencing, a map format and database used in the US

TNC Transportation Network Companies, including Uber and Lyft

TPM transition probability matrix in the Markov process

UE user equilibrium

USGS United States Geological Survey

VHT vehicle hours travelled

VKMT vehicle kilometres travelled

VISSIM a commercial microscopic traffic simulation programme

WLPPM Whole Life Performance Prediction Model

WSP a Canadian management and consulting firm

Chapter 1

Introduction

The transport sector is a prominent contributor to the greenhouse gas (GHG) emissions, estimated to constitute 14% of the global sum in 2010 and 28% of the US inventory in 2016 [9, 71, 112]. The question of how to reduce the transportation CO_2 emissions has been a primary concern for many transportation agencies across the world. In many aspects, innovations and conscious efforts to reduce the transportation system carbon emissions have been actively explored, which have brought a great amount of insights in fields such as traffic control, asset management, alternative energy and so on [29, 31, 145, 240, 248].

One challenge when extending the field-specific knowledge to the real system is to account for the multiple players operating on the same canvas. First of all, there are interactions to consider. For example, a unilaterally designed eco-routing system may lead vehicles to pothole-riddled street that ends up with higher emissions and vehicle operating costs [45]. Also, without a systematic perspective, it is often unclear how significant savings from one particular measure are in a broader context, such as the emissions due to traffic delays at pavement rehabilitation sites.

So far, most of the emission mitigation measures for the transportation system have been studied in silo, on project level or by reductionist approaches. Fortunately, with opportunities brought by the massive amount of open data in transportation, it is now feasible to scale previous analyses which rely on meticulously curated specific datasets to the system level. In this thesis, one vehicle mobility model and one pavement degradation model are developed in this manner with the aid of city-wide open data for the city of San Francisco. The vehicle mobility and pavement degradation processes, representing two most visible strands in the transport system, are brought together for the simulation of the carbon emissions in a full urban road network. The contribution of this work is (1) to develop efficient city-scale models that can simultaneously evaluating multifaceted carbon mitigation policy analysis

on one integrated platform; and (2) to quantify the strength of interactions and values of the combined carbon mitigation measures.

1.1 Background

1.1.1 The traffic-related carbon emissions and mitigation measures

Globally, road transport represents over 70% share in the total fuel consumption and carbon emissions in the transportation sector, compared to air or rail transport [113]. As a result, many options have been proposed to decarbonise the transportation sector from the road transport aspect [29, 107]. These options often include:

- Improving vehicle fuel economy.
- Developing low carbon alternative fuels.
- Reducing travel demand through modernising public transport, tele-working, etc.
- Increasing network efficiency with intelligent transportation systems (ITS).

These approaches are often the basis for estimating changes in transport fuel use or emissions in the future. For example, United States Energy Information Administration (EIA) projects the fuel economy of light-duty vehicles to increase by more than 40% from 2015 to 2025 with technology improvements and policy standards, even offsetting the forecast increase in travel demand in terms of vehicle-miles travelled in the same period [70]. However, such predictions are based on a series assumptions about technology, market acceptance rates, financial situations and fuel prices, which may lead to big variabilities and substantially different outcomes than reality as indicated by the past experience [165].

ITS has been identified as a promising niche innovation in the transition towards a low carbon transport system [88]. The central promise of ITS innovations are the enhancement of system efficiency with information, communication, control, computer technology and other current technologies [5]. The benefits of ITS include improvements in terms of safety, reduced congestion as well as sustainability impacts. The carbon mitigation potential of ITS applications have been demonstrated on intersection level (adaptive traffic signalling, self-coordinated virtual intersection control), link or corridor level (incident management, cooperative adaptive cruise control) and network level (eco-routing) [29, 81, 149, 230].

However, current evaluations of ITS application scenarios have two major drawbacks. The first is that many studies are reported at a single intersection or for a single vehicle that may not generalise to the real city-scale situations. Secondly, despite the great advancements in

traffic technologies, the transportation infrastructures are ageing quickly and can cast negative impacts on the total carbon emissions due to maintenance needs and increased operational costs. However, the impacts of ageing infrastructures are not considered in the evaluation of most ITS technologies. This thesis aims to conduct a city-scale evaluation of transportation carbon mitigation measures by incorporating the contributions from two sides, the ITS information that brings about changes in traffic patterns, as well as the contributions from the infrastructure asset maintenance, specifically road pavement maintenance scheduling.

1.1.2 The pavement-related carbon emissions and mitigation measures

Similar to the decarbonisation efforts from the traffic operation perspective, the CO_2 mitigation issue is also attracting awareness among the infrastructure asset managers represented by the pavement engineers. CO_2 emissions related to the pavement infrastructures are often assessed through Life Cycle Analysis (LCA) by considering the emissions in the following five phases: raw materials production, construction, use, maintenance/rehabilitation (M&R) and end-of-life dismantle/recycle [205]. Similarly, in the European Standard *EN 15643-5:2017* prepared by the Comité Européen de Normalisation (CEN) technical committee 350, pavement, as a form of the civil works, has the following life cycle stages: pre-construction (A0), product (A1-3), construction process (A4-5), use (B1-8), end-of-life (C1-4) and those beyond the system boundary (D) [75]. Opportunities for emission reduction arise in various aspects in the pavement life cycle components, as indicated by existing research [47, 240, 247].

Components associated with the greatest energy use and GHG emissions are: material production, use phase and delay/congestions during the construction or M&R. For the material component, large amounts of asphalt concrete and Portland cement concrete are typical used for new pavement constructions (rare in many urban areas) or rehabilitation (replacing the existing surface with new materials). The production of cement or asphalt emits GHGs including CO_2, CH_4 and N_2O, with the quantities of CO_2 dominating by three to six orders larger over other GHGs [249]. The proportions of material-related carbon emissions over the whole life cycle vary depending on the assumptions used in each study (e.g., the traffic level, materials used). For the a fairly busy road with Annual Average Daily Traffic (AADT) of 70,000 vehicles (8% trucks), the material component accounts for 30%, 14% or 11% of the total life cycle CO_2 emissions for pavements with surface type of Portland cement concrete, hot mixed asphalt or composite materials [249].

The CO_2 emissions from the pavement use phase encompass a variety of items: vehicle energy loss to overcome rolling resistance, albedo (reflection of solar radiation), carbonation (for Portland cement concrete), night lighting and leachate [206]. The use phase emissions have been shown to be a significant component in pavement LCA, even though this phase is

frequently ignored in studies due to the complexity [93, 106, 159, 178]. Like the materials, contributions of use phase emissions depend highly on the individual context of pavements. In a comprehensive sensitivity study, the global warming potentials (emissions of CO_2-equivalent) of eight LCA components are compared subjected to uncertainties in traffic volumes, pavement classes and other parameters [204]. They found that some components, such as carbonation, only have small contributions to the overall impact. While the rolling resistance component during the use phase becomes the dominant one in high traffic volume circumstances.

The use phase emissions are closely related to the pavement M&R, as well maintained roads can offer significant benefits in reducing the emissions from vehicle operations. Degraded pavements, usually characterised by a higher surface roughness caused by surface defects, will create more vibrations and rolling resistance for vehicles, thus increasing the fuel consumption and carbon emissions. A widely used indicator of pavement roughness is the International Roughness Index (IRI). It measures the cumulative suspension motion in a moving vehicle over the travelled distance [179]. An IRI of 0 m/km represents a perfectly smooth and flat road surface, while an IRI of around 10 m/km indicates severe defects. Airport runways usually maintain an IRI around 1 m/km while newly constructed urban roads have an IRI around 1 or 2 m/km. Previous studies suggest that a 1 m/km reduction in rolling resistance will lead to 1 or 2% savings in fuel consumption [240, 251]. The M&R works to restore pavement roughness induce additional energy use and emissions depending on the treatment type, but this extra cost can be offset in a short time by the savings of the subsequent use phase for high traffic volume roads and result in net reduction of the life cycle emissions [240, 238].

The last component with potential significant contribution to the reduction of life cycle emissions is related to easing the congestions or disruptions to traffic during the construction and M&R. This component captures the emissions from vehicles that are slowed down or detoured due to roadworks, but quantifications of its impacts from existing studies vary greatly. Generally speaking, the impacts of roadwork-related traffic delays depend on several factors, including the traffic volume, construction hour (peak vs off-peak), the loss of capacity as well as the availability of detours [204]. In Inti et al. [114], the GHG emissions due to the traffic delays of a 1.6 km six-lane highway (major rehabilitation, with partial lane closures, 130,000 vehicles per day) are found to be several times higher than the material component, while the use phase is not considered. In Huang et al. [108], the extra CO_2 emissions due to roadwork disruptions on a 2.6 km dual carriageway (full reconstruction, with partial lane closures over five days, 26,000 vehicles per day) are only about 1% of those from the material component and almost negligible compared to the usage phase. The extra CO_2 emissions

of 0.5% to 11% from traffic disruptions are reported on top of those from the rehabilitation process of a 0.7 km carriageway (major rehabilitation, with partial lane closure or diversion, about 10,000 vehicles per day) [83]. Emissions from roadwork-related traffic disruptions are not considered in some studies on the ground that M&R are carried out overnight in places such as California [138, 240].

Similar to the evaluation of traffic-related emissions, assessment of pavement CO_2 emissions also suffers from the problem of scale and the simplification of the complex traffic patterns. Many the studies reviewed are conducted based on one or several representative functional unit, usually one lane-mile of pavement section. However, in reality, the pavements in a transportation network have diverse characteristics in terms of material and construction quality, traffic loading, micro climate and maintenance history. Estimations of emissions based on one pavement section may not generalise to a network. In addition, the representations of the traffic patterns are usually simplified by assuming a fixed traffic level. The traffic redistributions due to increased user cost during the use phase or roadworks are not adequately addressed. These deficiencies point to the need of a more integrated perspective (e.g., coupling with traffic simulation) in the evaluation of the CO_2 emissions of the pavement infrastructures.

1.1.3 Interdisciplinary perspective: combining and comparing carbon mitigation options in traffic operation and infrastructure asset management

It is evident from the existing studies that both efficient traffic flow operations and well managed road infrastructure asset can contribute to the reduction of CO_2 emissions in the transportation system. However, carbon mitigation strategies from the two perspectives are seldomly compared or evaluated together. In fact, the infrastructure provisions are seldomly taken into account in the emission analysis or policies designed from the traffic perspective. Similarly, the changes in vehicular flow patterns devised by traffic engineers towards the low carbon emission goals are not reflected in the pavement management scenario analysis, either.

There are some literatures taking the interdisciplinary perspective to evaluate the carbon emissions of the transportation system. In Saxe et al. [207], a holistic framework is proposed that assesses the GHG emissions of the rail infrastructure projects in a complete manner, accounting for the capital GHGs, operational and maintenance emissions as well as the impacts of new infrastructures on ridership and urban forms. This framework is later applied to calculate the GHG emissions and payback time of the Sheppard Subway Line in Toronto,

Canada [208]. Chester and Horvath [46] proposed a comprehensive LCA that includes the infrastructure and supply chains into the assessment of passenger transportation. They found that by expanding the system boundary, the LCA energy use and emissions increase by an additional 40-70% for the road transportation system over vehicle tailpipe operation. Their study is based on breakdown of factors to normalised analysis units (cost per passenger-kilometre-travelled for each manufacture procedure, vehicle operation mode, etc.) rather than a full-scale simulation. Given the complexity of the problem, the breaking down of individual factors facilitates the quantification and comparison of contributions from each component systematically. However, it is still difficult to model the interactions between different components. For example, maintenance is scheduled every 10 years regardless of the type, traffic, condition of the pavements or budget constraints. Vehicle fuel economy is also assumed to be independent of the change in pavement roughness. In another study [230], the authors compared the energy and environmental benefits of an ITS based highway incident management system against four conventional construction phase carbon mitigation strategies (using regionally provided materials, reducing fossil fuels, recycling pavements and using warm-mix asphalt). They found that the long-term energy and environmental savings will far surpass the one-off savings from the construction phase strategies. Although not including the long-term benefits of well-constructed and well-maintained pavements, their study certainly highlights the uneven contributions from different carbon mitigation strategies.

1.2 Research objectives

The overarching objective of this research is to compare the effects of carbon emission mitigation measures from the traffic operation as well as the infrastructure management perspectives by building the traffic simulation and the infrastructure degradation prediction modules into one unified computational framework. These two perspectives are both crucial, and possibly interconnected, in reaching the low carbon emission goals. However, as introduced in Section 1.1, they are seldomly considered together for emission analysis. As a result, a necessary step must be taken first that ensures suitable and compatible methodologies are employed for developing the traffic simulation and infrastructure condition prediction modules. Based on this step, emissions and other network performance metrics can then be evaluated. The overall research objective breaks down to:

1. Development of models that are efficient for large scale analysis, complex for representing realistic behaviours and straightforward for system-level integration.

(a) Hybrid mesoscopic traffic simulations that combine the macroscopic traffic behaviours with disaggregated microscopic travel demand have been known for its efficiency, fine granularity and scalability to a large network [40]. It is adopted in this research with inputs and validations based on open data.

(b) A spatially autocorrelated link-level pavement degradation model is also adopted for its ability to represent the spatial variability and spatial structures in the pavement degradation process, as well as the compatibility with the link-based mesoscopic traffic simulations. Unlike the simulation-based traffic model, the pavement degradation model is statistical in nature. Model coefficients are learnt from the open infrastructure data.

(c) Identification and quantification of interactions between the traffic movements and the infrastructure conditions.

2. Application of the developed traffic and pavement degradation model to the study area to calculate the carbon emissions under different mitigation measures. Specifically including

   (a) Determination of the carbon emission calculation procedure that takes into consideration the traffic as well as the pavement conditions.

   (b) Quantification of the network performance in terms of the CO_2 emissions, the traffic efficiency and infrastructure conditions under different short and long-term scenarios.

   (c) Evaluation of the individual and combined effects of eco-maintenance (road works prioritising CO_2 reduction) and eco-routing (route selection prioritising CO_2 reduction) at the network level.

It should be noted that it is not within the objectives of this thesis to conduct a LCA of the transportation system. Given the few new construction projects in big cities and the high levels of traffic, this thesis chooses to focus on the use phase emissions due to vehicle infrastructure interactions. Embodied carbon emissions in the road infrastructure, which refer to the emissions contained in the construction and maintenance materials or those produced due to construction or roadworks, are thus not considered. Nor are the CO_2 emissions related to the production, maintenance and disposal of vehicles included in the analysis. Some carbon mitigation scenarios involves adjusting the pavement maintenance scheduling to achieve CO_2 reductions. Emissions related to such activities are generally not considered as well, for it is assumed that the total length of road maintenance remains the same in different scenarios.

1.3 Conceptual framework

An integrated traffic simulation and pavement condition prediction model is needed to address the research objectives identified in Section 1.2. The conceptual framework of this integrated model is presented in Figure 1.1. Starting from the outside loop, the traffic simulation module and the pavement degradation module serve as the two pillars of the overall framework. The traffic simulation module is sensitive to the changing pavement conditions (outputs of the pavement degradation module) and produces individual vehicle routes and street-level traffic volumes. The output traffic volumes from the traffic simulation module then enter the pavement condition prediction module to determine the potential traffic-induced degradation as well as to select important roads that receive higher maintenance priority. Together, the traffic volume outputs from the traffic simulation module and the pavement condition outputs from the pavement degradation module are combined to calculate the CO_2 emissions due to traffic usage and pavement roughness. The effects of different carbon emission mitigation measures are also incorporated. For example, by setting a new maintenance strategy, roads selected for maintenance will be changed. While the vehicle operation strategy affects the route choices of individual vehicles in the traffic simulation. Scenarios with different combinations of maintenance and vehicle routing strategies are compared in terms of the traffic efficiency, the infrastructure conditions and the CO_2 emissions, all at the network scale.

1.4 Thesis layout

Having presented the background, the objectives and the conceptual framework of this study in Chapter 1, the rest of the thesis is organised as the following:

A literature review of relevant topics is conducted in Chapter 2. It consists of three parts. The first part briefly reviews the concepts of sustainability, sustainable transportation and the value of models in the derivation of sustainability indicators. Although this research is not a comprehensive sustainability analysis, the outcomes (network-level CO_2 emissions, traffic efficiency and infrastructure conditions) can be beneficial in future sustainability studies. Then in the second part, the concepts for traffic modelling are reviewed, together with established methodologies for demand generation and supply simulation. In the third part, existing pavement degradation models are reviewed, with special focus on the commonly adopted model forms, inputs and outputs. The aim of the literature review is to identify suitable model classes and theories to be adopted in the subsequent chapters of this study.

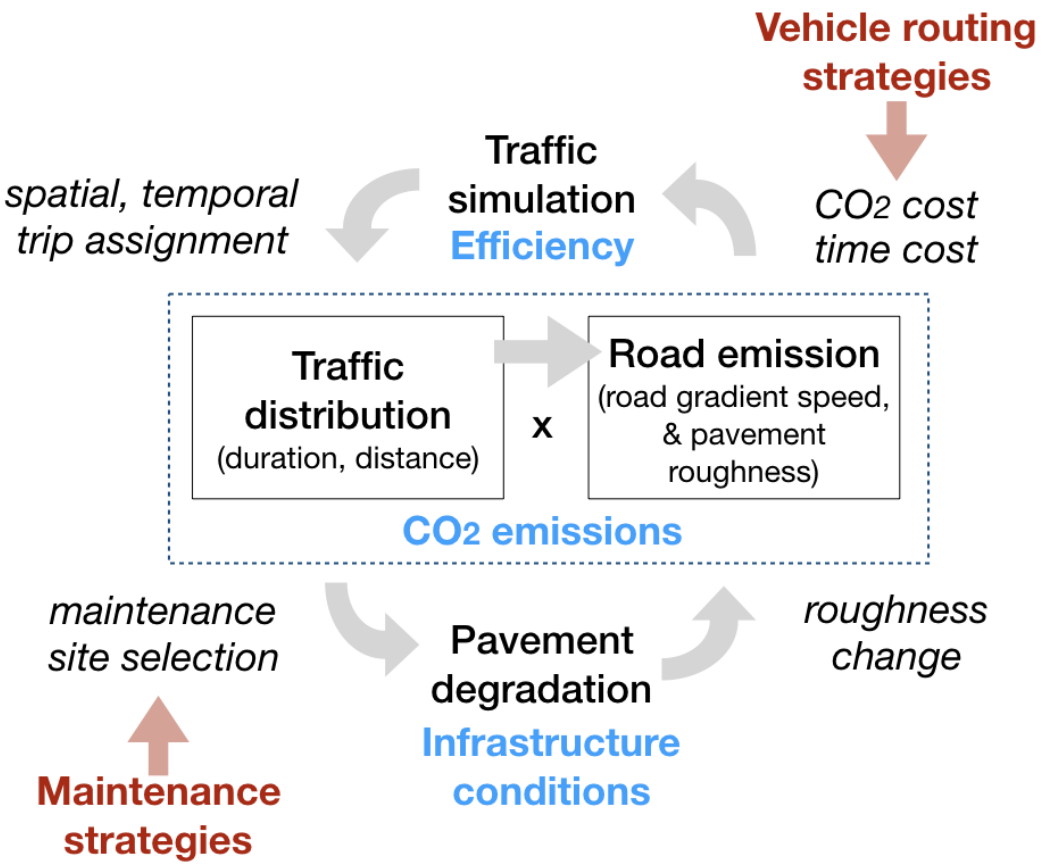


Fig. 1.1 Conceptual framework integrating the traffic simulation and pavement degradation functionalities. The outputs of the integrated model include traffic efficiency measures, infrastructure conditions and CO_2 emissions, all at the network scale.

In Chapter 3, the development of a mesoscopic traffic simulation model is presented for the case study area of San Francisco. Detailed data collection and cleaning procedures are documented to ensure reproducibility. The functionality and performance of the developed traffic model are demonstrated.

The validity of some underlying assumptions in the mesoscopic traffic model presented in Chapter 3 is further investigated in Chapter 4. Specifically, travellers do not have full access to the real-time traffic condition information and such information can only be revealed by specially dedicated probe vehicles. In addition, random repetitions are carried out to quantify the variability of the results due to the randomness embedded in the traffic assignment procedure. The aim is to quantify the impacts of information availability and quality on the outcomes of the proposed mesoscopic traffic simulation tool.

Chapter 5 presents the process to build city-scale pavement condition prediction model with more than two decades of pavement condition survey data of the case study area. Two

degradation predictors (pavement age since last maintenance and cumulative traffic load) and three model structures (non-spatial road category-based model, non-spatial individual street model and a spatial hierarchical model) are tested. The performance of the candidate models are evaluated in terms of their accuracy and complexity. The spatial model with pavement age as the degradation predictor is found to have the best performance.

In Chapter 6, the traffic simulation model (developed in Chapter 3) and the spatial pavement degradation model (developed in Chapter 5) are integrated jointly to simulate various CO_2 mitigation scenarios of the study area. The individual and combined effects of the eco-routing traffic operation strategy and eco-maintenance for pavement asset management are evaluated in terms of several network-wide metrics, including the Annual Average Daily vehicular CO_2 emissions, Annual Average Daily vehicle hour/distance travelled and the overall pavement conditions over an analysis period of 10 years. Complex factors such as traffic growth and fuel types are not considered in most parts of Chapter 6 to exclude their interferences in result interpretations. However, additional results based on traffic growth scenarios are given at the end of Chapter 6 to demonstrate the sensitivity of the outcomes.

Lastly, the summaries of findings in each chapter and the conclusions are presented in Chapter 7. The research objectives introduced in Chapter 1 are also revisited in Chapter 7 to assess whether they have been met through the analyses.

Chapter 2

Literature Review

Following the introduction of the research objectives to assess the CO_2 emissions and other traffic and infrastructure performance metrics of a city-scale road network, relevant literatures are reviewed in this chapter. First of all, the concepts of sustainability and sustainable transportation are introduced. Although it is not the aim of this thesis to conduct a complete sustainability analysis, the various quantities of interest (e.g., network-level CO_2 emissions, traffic efficiency and infrastructure condition metrics) correspond to certain frequently used sustainability indicators and may be adapted for sustainability studies in the future. On these network-level outcomes, traffic conditions and pavement conditions are the two major aspects that have potentially significant impacts. Existing literatures are reviewed for methodologies to calculate/predict the traffic and pavement conditions. When there is a need to incorporate future scenarios, model-based estimations are still the most prevalent methodologies adopted in existing literatures. Highly detailed models usually require more complex data inputs and are more computationally extensive, while simple models that aim to capture key trends have the advantage of solving larger scale problems efficiently. To guide the selection of the modelling methodologies for the later part of this study, special attention is paid to the complexity and input/output requirements of different models in this review.

2.1 Sustainability concepts and sustainable transportation

The dialectic contradiction of resource conservation and economic growth has long been co-existing. For example, since ancient times, crop rotation and rest-rotation grazing were adopted to replenish soil nutrients or forage supplies [82]. Even though some pre-industrialised societies also faced ecological problems such as deforestation or land exhaustion, the tension between conservation and growth became acute after the unprecedented level of resource exploitation and population growth after the industrial revolution [39]. A series

of events in the twentieth century, notably the Dust Bowl in North America in the 1930s, the publication of the environmental science book *Silent Spring* in 1962 and the modern environmental movements in the 1970s, led the world to reflect on our relationship with the nature and the future [105, 166]. It was under this background that the concept of sustainable development was put forward in the 1970s [153]. In 1987, the United Nations World Commission on Environment and Development released the Brundtland Report, formally titled *Our Common Future*, which produced the most cited definition of sustainable development till today [32]:

"Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs."

The ideal of sustainable development also appeals to the transportation sector, where the trade-offs between the increased level of mobility and the reduction of congestion, pollution and safety issues need to be reconciled [142]. Transportation is a driving force for economic development and poverty alleviation, while in the meantime, a significant contributor to the greenhouse gases (GHGs) and air pollutants alike [237, 245]. In addition, the Word Bank estimates 1.3 million deaths each year in traffic-related accidents [214]. In a review of sixteen national and international sustainable transportation initiatives, the definitions of sustainable transportation are found to be organisation specific, but nonetheless are built along three dimensions: economic development, environmental preservation and social development [119]. These three aspects are adopted in the European Standard on the sustainability assessment of buildings and civil engineering works [75] and are sometimes called the "triple bottom lines", a term borrowed from business accounting [72]. Indicators belonging to these three dimensions are often adopted for assessing the sustainability of a transport system or project, such as the vehicle kilometres travelled (VKMT, economic dimension), per capita traffic cashes and fatalities (social dimension), per capita air pollution emissions (environmental dimension), etc. [141]. Given the prolific numbers of indicators, frameworks have also been developed for selecting and combining different indicators based on criteria such as measurability, availability, interpretability as well as their relevance to the sustainability goals [42, 194].

In many transportation sustainability studies, data used for calculating sustainability indicators are directly observed or measured [255]. In Reisi et al. [194], final indicators (e.g., energy use and GHG emissions) are deduced from measurable quantities (e.g., VKMT) based on simple models, which provide the flexibility and sensitivity required for policy design. Going beyond the need to assess the sustainability of the current system, model-based indicator quantification then becomes essential for analysing future scenarios [117, 134].

In this thesis, carbon mitigation strategies from two perspectives (traffic operation and infrastructure maintenance) are compared. While this is not a fully fledged sustainability study, some evaluation metrics, namely the network-wide CO_2 emissions, VKMT, infrastructure conditions, coincide with the sustainability indicators. The analysis in this thesis can potentially contributes to the more holistic transport sustainability studies in the future. The rest of this chapter serves to review the concepts and methodologies related to the modelling and calculation of the evaluation metrics.

2.2 Traffic Models

For short-term traffic condition predictions, assumptions of unchanging states or statistical methods that extrapolate the spatio-temporal trends based on real observations are often used. For example, an eco-routing system is presented in Zeng et al. [252]. The system is built with real traffic condition data collected by 153 probe vehicles. It took 10 months to collect the average link-level speed and acceleration data in the case study area of Toyota city, Japan (4,072 nodes and 12,877 links, with an average of 11 observations per link). These data, combined with the road gradient, are used to determine the link-level emissions. Such real observation based traffic condition forecast can lead to accurate short-term predictions. However, the method cannot be extended to predict unforeseeable scenarios, where the traffic distribution characteristics might be considerably different from the existing observations. As a result, for more versatile scenario analyses, simulation based traffic models are still the preferred choices.

There are many classification schemes that have been used to categorise the traffic simulation models. For example, the macroscopic and microscopic distinctions, the static and dynamic traffic assignments, the trip-based versus activity-based models, etc. In Nagel et al. [162], the authors list the following components as "absolutely necessary" for traffic planning models: (1) demand generation, (2) supply simulation, (3) feedback and (4) analysis. According to Nagel et al. [162], traffic dynamics are created when the demand to travel is constrained by the network supply. Iterations are carried out in such simulations to make the demand and supply consistent (feedback), while analyses are needed to obtain further insights (e.g., congestion distributions). This high-level summary is found to offer the clearest guidance for the review of the vast traffic simulation literature and is adopted in the organisation of this section, with focus on the demand and supply modelling.

2.2.1 Demand generation: trip-based and activity-based models

There are three types of commonly used demand generation models: the simple trip-based (four-step) models, tour-based models and the most sophisticated activity-based models. The classical and still most widely adopted transportation planning models are trip-based. They are also called four-step models due to the four distinct modelling steps involved [150]:

1. Trip generation;
2. Trip distribution;
3. Mode choice;
4. Route choice.

Trip-based models often operate on aggregated spatial units, such as census tracks or Traffic Analysis Zones (TAZ). In the first step (trip generation), land use models or zonal socio-demographic statistics are used to determine the numbers of trip origins and destinations associated with each aggregated spatial unit. In the second step (trip distribution), zonal origins will be matched to destinations in other zones to form an origin-destination (OD) matrix. Individual OD matrices may be generated for each time period of a day, such as the morning peak, the off-peak and the evening peak periods. In this sense, each element in the OD matrix represents the travel demand from one origin zone to one destination zone and is aggregated both spatially (into zones) and temporally (into time periods or a whole day). The travel demand given by the OD matrices can then be assigned with different modes of travel, e.g., by passenger cars or public transits in the third step (mode choice). Then, the fourth step (route choice) is carried out to assign the optimum path based on criteria such as distance, expected or experienced travel time, journey cost, travel time reliability, etc.

One of the earliest and also the most influential implementations of the four-step model is the Chicago Area Transportation Study (CATS) in the late 1950s and early 1960s under the leadership of J. Douglas Carroll [21]. In CATS, the four-step model was used to estimate traffic volumes on each street within the study area. In particular, since the route choice (sometimes called trip assignment) was computationally heavy, the study prompted one of the earliest applications of computers in traffic modelling. Nowadays, the calculation of route choices on a large network is still expensive, as encountered in this thesis and other studies [89]. Although many limitations have now been realised regarding this trip-based modelling approach, particularly regarding its aggregated spatio-temporal representation and inflexible demand specification, it is still a landmark technique in traffic modelling and a valid approach for planning for urban transportation as a system, rather than focusing on individual streets or corridors.

One criticism of the trip-based models is its treatment of trips made by the same person as independent decisions [197]. This is not behaviourally sound given many of the trips are actually "chained", i.e., home-work trip followed by a work-home trip. To improve the four-step models by considering the sequences of trips, tour-based or activity-based models are proposed. The tour-based model is considered as an intermediary step between the four-step model and the fully activity-based demand generation approach. The unit of travel in tour-based model is defined as a sequence of trips from home to several destinations and then back home [197]. In Krizek [131], a study is presented regarding the relationship between the socio-economic characteristics (household and neighbourhood characteristics) and travel behaviours at the tour level (type, frequency and distance). In the literature, the tour-based model is particularly popular for the modelling of commercial vehicles (deliveries or other reimbursed trips made during working hours), whose travel itineraries can be derived from the supply chains [110, 122].

Activity-based models offer even greater flexibility to model the complex travel demand generation process than the tour-based approach. In activity-based models, the unit of travel is an entire day or entire week's schedule of each individual that consists of multiple tours [30]. Travel demand is derived from the need of individuals to carry out certain activities. A traveller may rearrange his/her schedule, add or cancel some trips by evaluating the net utilities of the whole travel plan. In Zheng et al. [256], the activity-based traffic modelling is applied to a dynamic cordon-based congestion pricing scheme. In this application, the change in travel cost affects not only the route choices, but also the daily plan of the travellers. The pricing scheme leads to about 5-20% of travellers to shift their departure times away from the congested periods and results in the 20% higher econometric values of travel time savings than the toll costs.

In Hatzopoulou and Miller [102], traffic conditions generated by activity-based demand models are used to calculate traffic emissions. The authors conclude that using activity-based demand generation enables the consideration of finer time steps than traditional approaches. Also, the activity-based demand generation allows the tracking of individual travellers for calculating person-level air pollutant exposures. However, unlike the cordon pricing study in Zheng et al. [256], no feedback mechanism is implemented to modify the activity patterns based on the pollution exposures in Hatzopoulou and Miller [102]. As a result, it is still not very convincing as for the necessity of adopting the activity-based approach for their emission analysis, i.e., why traffic generated by a simpler trip-based models (with finer spatio-temporal disaggregation) would not yield similar results.

Activity-based models are good at representing realistic travel behaviour by expanding the four-step trip-based approach, which is instead more statistically-oriented and assumes

Table 2.1 Travel demand generation models

Model type	Trip-based	Tour-based	Activity-based
Unit of analysis	Single trip (e.g., home-work)	A sequence of trips (e.g., home-work-shop-home)	Daily or weekly schedule of an individual
Spatial resolution	Usually zones (e.g., TAZ)	Household, individual	
Temporal dynamics	Fixed for long time periods (e.g., morning peak)	Choice of travel time	
Inputs	Basic demographic inputs (e.g., household size, income class)	Detailed demographic inputs	
Computation effort	Low	Medium	High

fixed travel schedules. However, the flexibility in activity-based models comes at the price of the complexity. The inclusion of all possible combinations of trip sequence, departure time, destination, mode and route choice to be explored would quickly explode the solution space [151]. Another major factor that limits the use of activity-based models is the availability of data. Travel surveys as the primary source of knowledge of regional travel demand are still largely trip-based. For surveys that focus on household activities, little has been collected regarding the spatial and temporal constraints of activity schedules [151]. In Griesenbeck and Garry [98], a summary is given based on the experience of using SACMET (trip-based) and SACSIM (activity-based) travel demand models from the perspective the regional planning agency in Sacramento, California. The trip-based SACMET requires less inputs (1528 TAZs, median size is 390 acres), data processing efforts (less than 5 days) and model running time (6 hours). While the more sophisticated activity-based SACSIM requires inputs from 650,000 parcels (average size is 0.8 acres), two-weeks for data processing and 24-30 hours of run time depending on the population simulated.

Facing the trade-offs of between complexity and fidelity, the selection of trip-based or activity-based demand model ultimately depends on the need, the data and the time resources available [257]. In real practices, more advanced models are favoured by jurisdictions with complicated traffic and congestion management scenarios. In such cases, activity-based

models are adopted to test a wide range of policy scenarios (e.g., congestion pricing, land use). Smaller agencies may find themselves unable to develop activity-based models due to the lack of resources and support, in which case the trip-based demand models are more often adopted for their simplicity [62].

2.2.2 Supply simulation: flow propagation models

The previous section describes the modelling methodologies to estimate the travel demand. Travel demand models are often paired with supply simulations which depict the abilities or characteristics of the traffic network in the handling of such demand. The supply of a road link or a traffic network is not constant due to vehicle interactions: when the number of vehicles on the road exceeds a certain threshold, the travel speed reduces and the flow (traffic volume in a unit time) also drops. The supply simulation is closely related to the modelling of traffic flow propagation and can be carried out at three levels: macroscopic, microscopic and mesoscopic levels [146].

Macroscopic traffic propagation models

Traffic flow propagation is often analogised to fluid flow and modelled as continuous compressible media. Similar to fluid dynamics, traffic propagation can also be described by speed, density and flow in the realm of macroscopic models. These parameters used by macroscopic traffic propagation models are measured by commonly adopted traffic sensors, such as the loop detectors.

Speed v: Various representations exist of traffic speeds, such as the instantaneous speed, journey speed, space mean and time mean speed. These speed representations are adopted depending on the context. For example, the instantaneous speed is used for accident and detailed emission analysis [129, 191]. For macroscopic traffic propagation models, the frequently used representation is the space mean speed, defined by the total travel distance over the total travel time, or other similar definitions (Equation 2.1) [12, 101]. Vehicle speeds are most commonly obtained from loop detectors. They can be directly measured by dual loop detectors ("speed traps") or estimated from single loop detector measurements assuming constant vehicle length [243].

$$v_s = \frac{N \cdot D}{\sum_i t_i} \tag{2.1}$$

where v_s is the space mean speed of traffic stream, N is the number of vehicles. D is the measurement distance and t_i is the time taken for vehicle i to go through the measurement distance.

Flow q: Flow, or flow rate, measures the numbers of vehicles going pass a point location in a unit length of time (Equation 2.2) [101]. Flow data can be measured through manual counting, loop detectors, radar, television cameras, etc. For high-level planning purposes, the flow measurement time interval can be as long as the peak hour period or a whole day (e.g., Annual Average Daily Traffic, AADT). To capture highly dynamic traffic variations, however, shorter measurement intervals at every 15 minutes should be used [163]. It is possible to poll the flow data every second, but it is recommended to use at least a 10 minute interval to ensure the stability of the flow measurements [219].

$$q = \frac{N}{T} \tag{2.2}$$

where q is the flow rate (vehicles per hour), N is the number of vehicles and T is the elapsed time.

Density k: density is the number of vehicles per unit length. It is widely used in the macroscopic traffic propagation models due to its theoretical link with the space mean speed and flow and as an indicator of the Level of Service (LOS) of the roadways. Direct measurements of density can be obtained from the aerial photography [177]. However, due to the high costs and difficulty associated with direct density measurements, they are often obtained indirectly from the flow and speed data based on Equation 2.3 under the assumption of homogeneous traffic conditions [9].

$$k = \frac{q}{v_s} \tag{2.3}$$

where k is the density (e.g., vehicles per lane per km), v_s and q are as defined above.

Macroscopic models describe the collective vehicle dynamics in terms of the average quantities (density k, flow rate q and space mean speed v) [76]. Any one of the three fundamental quantities k, q and v can be calculated from Equation 2.3 given the other two. Theories of macroscopic flow models to solve for the unknown quantities have been established in the seminal works of Aw and Rascle [11], Lighthill and Whitham [139], Newell [167], Payne [185], Richards [195], Zhang [254]. First-order macroscopic models (e.g., the celebrated Lighhill-Whitham-Richards (LWR) model) is based on vehicle conservation as a first order partial differential equation (Equation 2.5) and assumes a fundamental diagram

relationship (Equation 2.4). These two equations, combined with Equation 2.3, are used to solve for the fundamental traffic quantities. Solutions can be obtained by using analytical approaches such as the method of characteristics in simple settings, or numerically in other cases.

$$\begin{aligned} q &= q(k) \\ \text{or} \quad v &= v(k) \end{aligned} \tag{2.4}$$

$$\frac{\partial q}{\partial x} = -\frac{\partial k}{\partial t} \left(= \frac{\partial^2 N(x,t)}{\partial x \partial t}\right) \tag{2.5}$$

where $N(x,t)$ are the space (x) and time (t) dependent vehicle counts. q, k and v are flow, density and velocity as defined above.

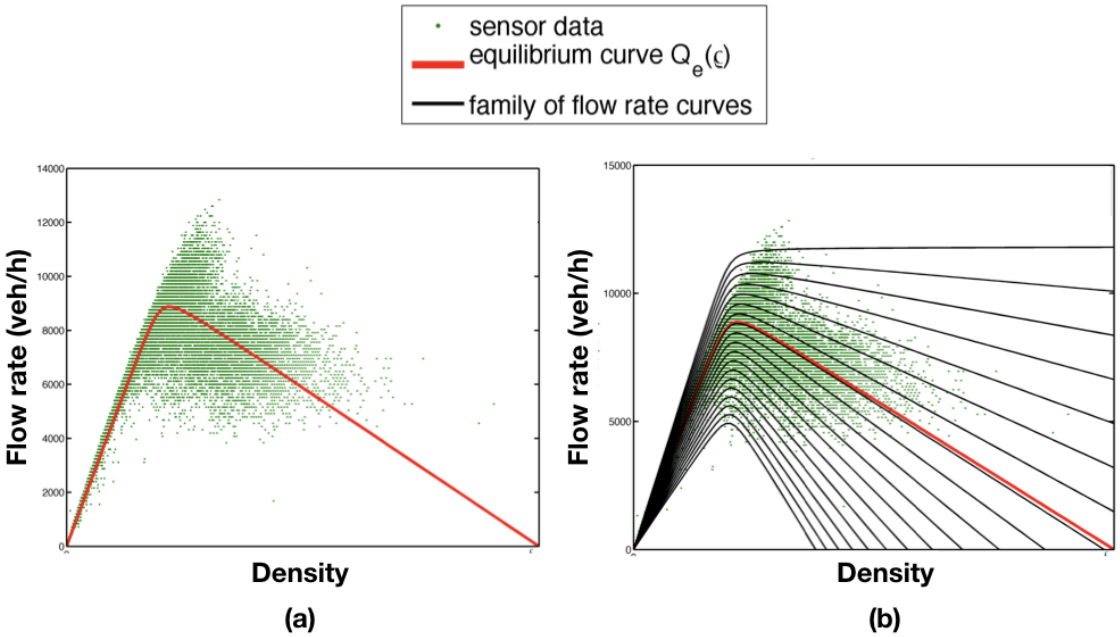


Fig. 2.1 First and second-order macroscopic models. (a) Fundamental diagram of the LWR model together with sensor data. (b) A family of flow vs. density curves of the ARZ model, with fundamental diagram and sensor data. Figures created by Fan [76].

First-order macroscopic models are conceptually simple and can capture important features of traffic flows, such as the shock waves (a short region of increased traffic concentration

where the speed changes suddenly). But they also have some shortcomings. For example, the flow-density relationship is represented by only one curve (the fundamental diagram), while in reality different classes of drivers may each hold a distinct fundamental flow-density relationship (Figure 2.1) [76]. Also, the fundamental diagram in Equation 2.4 of the first-order model implies that the speed is uniquely defined by the density, or in other words, vehicles have infinite accelerations when the density changes suddenly. This assumption is not very realistic, either. Higher-order macroscopic models have been developed to overcome these shortcomings of the first-order models. For example, in the second-order model proposed by Payne [185], the Equations 2.3 and 2.5 are preserved, but the fundamental diagram (Equation 2.4) is replaced by a momentum equation, linking acceleration (derived from v) with k (Equation 2.6).

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = \frac{V_e(k) - v}{T_{elp}} - \frac{c^2(k)}{k}\frac{\partial k}{\partial t} \tag{2.6}$$

where $V_e(k)$ is the "equilibrium" speed provided by the fundamental diagram in the first-order models, T_{elp} is the relaxation time to adjust the speed and c is an anticipation term as drivers anticipate on the downstream conditions (e.g., accidents) and adjust speed.

In the history of traffic propagation model development, an interesting debate occurred as for the significance of second-order macroscopic models. In 1995, Daganzo published a paper titled "Requiem for Second-Order Fluid Approximation of Traffic Flow", in which he pointed out the second-order model by Payne [185] is not anisotropic, i.e., vehicles may drive backward [61]. This is caused by the usage of the space derivative of density $\partial k/\partial t$ in the calculation of vehicle acceleration in Equation 2.6. When density increases rapidly along the space (i.e., $\partial k/\partial t$ is large), a negative speed can be predicted. To address this flaw, [11] proposed a new second-order model replacing the space derivative of density with a convective term that would eventually become known as the ARZ model.

Software tools and applications have also been developed based on the theories of macroscopic traffic models. One example is the METANET developed by Papageorgiou [176] incorporating second-order traffic flow models. In Kotsialos et al. [130], an application of METANET to the city-scale Amsterdam motorway network is presented (total road length is 143 km). The outcomes are validated quantitatively against the loop detector data at selected sections and qualitatively against the observed recurrent congestion patterns. Other examples include MARPLE and INDY, both of which assign travel demands to pre-generated routes [25]. Case studies of MARPLE and INDY are conducted on the Dutch national road network. This network contains 18,000 nodes and 25,000 links, which is at the same order

of magnitude compared to the case study network in this thesis. It takes 3 to 20 minutes to pre-generate around 150,000 routes on this network, while the actual dynamic traffic assignment of the 48,000 OD pairs (travel demand) takes 1 hour using MARPLE (30 second time step) or 5 hours on INDY (10 second time step). Overall, the numbers of variables in macroscopic models are independent of the numbers of vehicles, making them relatively easy to scale up to large networks [24].

Microscopic traffic propagation models

Microscopic models are characterised by the explicit modelling of vehicle interactions. Although there exist many theories regarding the handling of vehicle interactions, the best known one among them is the car-following model. As the name suggests, the car-following model specifies the reaction (e.g., acceleration) of the follower in response to the leading vehicle. The vehicle dynamics in the basic car following model are expressed by the following equation [146]:

$$\frac{dv_{follower}(t)}{dt} = \frac{v_{leader}(t) - v_{follower}(t)}{T} \tag{2.7}$$

where $v_{follower}$ and v_{leader} are the instantaneous speeds of the follower and the leader at time t. T is the relaxation parameter. Equation 2.7 states that the follower tries to match its own speed to that of the leader. Other variations of this basic form are available, such as including a reaction time τ at the left hand side of Equation 2.7, or the tendency to accelerate and maintain a desired speed or safe distance.

In microscopic models, it is convenient to handle lateral vehicle movements, such as lane-changing in a multi-lane setting. Lateral vehicle movements occur when a vehicle overtakes a slower vehicle, turns at an intersection or goes off a highway ramp. They may create disturbances to the stable traffic flow and impair the safety and capacity of the roads [126]. The lateral movements may also affect the following vehicles when they have to yield. Unlike the car-following model which can be modelled with a set of equations, the lane-changing behaviours are usually described by a set of criteria. In one of the earliest works modelling lane-changing, a structure is proposed to ensure that the simulated drivers make logical decisions by considering several factors, including the gap size and vehicle speed on the desired lane, the proximity of heavy vehicles, obstructions (e.g., parking area), etc. [92]. This has evolved to the distinction of mandatory, discretion and random lane-changing behaviours, where each is triggered by different motivations [190].

Due to the computational costs, microscopic models are often used in simulating small areas, such as a road intersection, a stretch of the highway or a ramp. Many commercial traffic

simulation software programmes are based on microscopic traffic propagation theories, with the most successful ones being the VISSIM and AIMSUN (which also has the macroscopic simulation ability) [3, 187]. These software tools usually offer graphical user interfaces that produce animated scenes of moving vehicles. An open source alternative for these commercial software programmes is the SUMO (Simulation of Urban MObility) [16]. SUMO implements car-following and lane-changing behaviours. Vehicle positions are updated at every second interval. A study applies SUMO for simulating vehicle platooning, where vehicles are tightly spaced so as to increase the highway capacity [78]. Due to the reduced spacings between vehicles in a platoon, the road capacity can increase by a few times than the free vehicle case. To test the proposed control and vehicle communication algorithms for platooning, simulation experiments are conducted on a one dimensional road link (5 km in length). The output variables of interest from the microscopic simulations include the velocity profile, as well as the distance and spacing errors of each individual vehicle to assess the formation process and the stability the platoons.

Mesoscopic traffic propagation models

In mesoscopic traffic propagation models, the notion of individual vehicles is still preserved, but the traffic characteristics are reported in aggregated terms, such as the space-mean average speed or the probability distribution of the headways. Three popular approaches are identified for mesoscopic traffic modelling in Maerivoet and De Moor [146]. The first approach is called the cluster models, where nearby individual vehicles are grouped into clusters and the flow characteristics are assumed to be constant within a cluster. The second approach is to specify the time headways of successive vehicles as a probability distribution, rather than modelled individually as in microscopic models. The third approach associated with mesoscopic traffic modelling is called the "gas-kinetic" models, where the traffic motions are in analogy to the motions of gas atoms or molecules and the vehicle density is described by a distribution at a given location x, time t and the space-mean speed v_s [233].

It is acknowledged among traffic modellers that the definition of mesoscopic models is somewhat ambiguous. Certain mesoscopic models are confused with microscopic models and are referred to as low fidelity "microsimulations" when individual vehicles are resolved [123]. In this review, models that retain individual vehicles but do not explicit express their interactions are categorised as mesoscopic, such as the queuing model or the cellular automata model. Some examples of mesoscopic models in the previous research are summarised below.

In the traffic network supply simulation of DynaMIT (Dynamic Network Assignment for Management of Information to Travelers) developed at the Massachusetts Institute of Technology, the microscopic representation of traffic is combined with macroscopic traffic

dynamic models. According to Ben-Akiva et al. [17], individual vehicles are retained in the simulation so as to model the impact of en-route information on driver decisions. While to be applied to real-time traffic operation management, the macroscopic traffic propagation models are adopted for the superior computational performance. The link or segment-level traffic dynamics are simulated with a queueing model by partitioning the road links into a moving part and a queueing part. The build-up of the queue is based on the position of the end of the queue at the previous time step, the vehicle length and counts, as well as the queue dissipation rate.

Another case of mesoscopic traffic model is the DYNASMART (DYnamic Network Assignment Simulation Model for Advanced Road Telematics) [118]. The macroscopic conservation law in Equation 2.5 and the fundamental diagram in Equation 2.4 are used in the traffic propagation calculation. However, unlike the macroscopic models, the flow q calculated by multiplying k and v in macroscopic models is replaced by tracking the local speed and position of each individual vehicle explicitly. The speeds of all vehicles in a link are the same as given by the second expression in the fundamental diagram (Equation 2.4). In the initial stage of the model development in the 1990s, DYNASMART could already handle the simulation of up to 75,000 vehicles and 2,000 links. But the most time-consuming step, the "path finding", is simplified. Specifically, the path finding algorithm is only executed at some intervals and a few, e.g., 5, paths are stored. Before the next round of path finding, the travellers can only choose from the existing paths, which might not be the absolutely optimum path. DYNASMART has been applied to many other applications. In Mahmassani and Abdelghany [147], a multi-objective assignment procedures are implemented in DYNASMART to model the traveller mode and route choices in an intermodal traffic network, including private cars, buses, metro lines and High Occupancy Vehicles (HOVs). DYNASMART is applied to the simulation of crowd evacuation in Kwon and Pitt [133]. Given a network consists of 2,488 nodes and 5,565 links and a evacuation demand of 30,000 to 40,000 vehicle trips, it took around 2 to 19 hours on a personal computer to simulate the evacuation process depending on the vehicle numbers and the complexity of the scenarios.

Table 2.2 Vehicle flow propagation models

Model type	Macroscopic	Microscopic	Mesoscopic
Conservation law	Partial differential equation of aggregated flow and density	Vehicles do not appear or disappear	
Vehicle motion	Fundamental diagrams relating space-mean speed with density	Individual vehicle longitudinal or lateral motions constrained by the motion of surrounding vehicles (e.g., speed difference, distance)	Fundamental diagram, or combining the the running part of a link, stop at the queuing part
Notable theories	LWR model (1st order), AR model (higher order), Payne model, Newell model	Car-following, lane-changing	Queuing theory, gas kinetics
Advantages	Efficient when scale up to a large network; the speed of solution does not depend on the numbers of vehicles; easy to calibrate with common traffic measurements	High-fidelity in modelling individual motion in complex scenarios (e.g., vehicle-vehicle interaction); handling lateral movements; individual motions can be animated	Handling individual vehicle routes and motions at higher efficiency than microscopic models
Disadvantages	No representation of individual vehicle trajectories	More computational intensive; hard to calibrate	Simplified or no vehicle interactions

MATSim: an example of a popular mesoscopic agent-based traffic model

Among the many traffic simulation software programmes available, the mesoscopic model MATSim-T is reviewed in detail due to its popularity, flexible agent-based framework as well as the successful applications in modelling large networks [43]. MATSim-T (Multi-Agent Transport Simulation Toolkit) is an agent-based traffic simulation software tool, where each vehicle is explicitly modelled as "agents". It uses a queue model to move agents between links. Figure 2.2 provides an illustration of the queue-based model concept. For each link in the network, some attributes are stored, such as the length of the link, free-flow speed, capacity and numbers of lanes. A vehicle entering a link (e.g, Road A) has to spend a minimum length of time t^A_{min} that equals to the free-flow travel time. The vehicle can then move to the next link if the capacity of the current link has not been exhausted, as well as there is space in the next link. If any of these conditions cannot be met, the vehicles have to queue at its current link. The queue-based traffic simulation model is capable of simulating the spill back of the queue to the upstream links. But it is not able to correctly represent sub-link dynamics such as the backwards travelling "kinematic waves (shock waves)".

MATSim-T uses two loops to implement the queue-based model. A buffer is added to each link holding all the vehicles that can leave the current link. In the first stage, all links are looped through to move any eligible vehicle from the current link to the end buffer, if these vehicles have spent the minimum time and the total outflow is below the capacity of the current link. In the second stage, the algorithm loops through each intersection instead and moves vehicles from the upstream buffer to the empty space on next link. If there is more than one upstream buffer and the downstream space is limited, vehicles in the buffers are served (moved to the next link) proportional to the link capacity.

The performance of MATSim-T, as well as other simulation-based traffic models, is largely influenced by how often the network states (link volumes, vehicle positions, ...) are updated. In MATSim-T, two updating approaches have been implemented. The first one performs a straightforward time-based updating. Movements of vehicles and the in-flow and out-flow of each link are calculated at one second intervals (time in the simulation, not in the real life). The overall computation time thus depends on the number of road links to be updated in the network. Another approach is to update the system states only when "event" occurs, such as a car entering or leaving a link. In this case, the computational time is proportional to the number of vehicle agents and the simulation speed is significantly faster when the network is empty (e.g., in off-peak hours). According to Charypar et al. [44], even in the simulation of rush hour traffic, the event-based approach still outperforms the time discretisation method.

The queue-based traffic simulation model in MATSim-T is paired with the activity-based demand generation (as opposed to the traditional four-step models that produce fixed agent schedules) to optimise agent departure times and route choices. This leads to a steady state approximation of the dynamic Nash equilibrium [13]. Other applications include the estimation of trip externalities (emissions, noises, congestions, etc.) [227] and evaluating congestion charging effects [256].

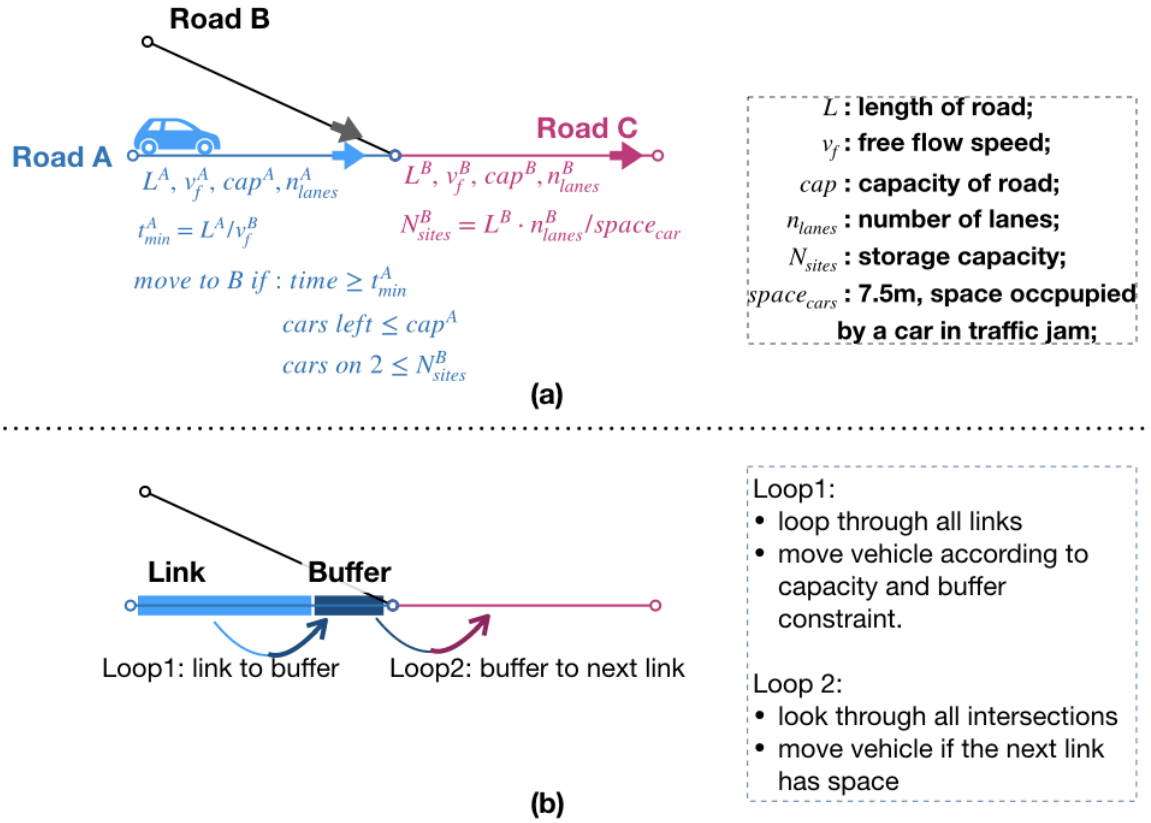


Fig. 2.2 An illustration of the queue-based model in MATSim. (a) Vehicle motions in the queue-based model; (b) Buffer.

2.2.3 Summary of the review of traffic models

In the above sections, some of the key theories related to the modelling of the traffic flow are reviewed. Specifically, the traffic flow is regarded as applying the travel demand to the road network, which has limitations in the supply. Travel demand can be generated by trip-based, tour-based or activity-based models. As the travel demand is modelled in an increasingly realistic manner, e.g., by considering the sequence of activities that an individual performs, the complexity of the demand generation also increases.

The supply side of the traffic network determines the traffic flow characteristics under the given demand and can be modelled at different levels of fidelity. In the most aggregated form, macroscopic traffic propagation models involve specifying the relationships between macroscopic traffic measures, such as the flow-density or speed-density relationships. The microscopic models, on the contrary, explicitly model the motions of individual vehicles and their interactions by implementing car-following or lane-changing logics (related to the longitudinal and lateral movements of vehicles, respectively). The third type of traffic propagation model adopts a mesoscopic approach, which retains individual vehicles but models their movements according to the macroscopic traffic propagation theories. Given the efficiency and disaggregated demand representation, the mesoscopic models are especially suitable for analysing large-scale scenarios while considering individual traveller's decision making process.

2.3 Pavement Degradation Models

The pavement degradation models surveyed span over a great variety, from simple deterministic models to complex stochastic ones. Some are at a conceptual level (e.g., the detailed pavement-vehicle interaction model proposed by Collop and Cebon [53]), while others are already implemented in specific countries, regions, or even worldwide [7, 124]. The models usually have a long full name and a short acronym, both of which are provided in Table 2.3. Only the abbreviations are used in the following review.

2.3.1 Nature of the modelling methodology

Based on the nature of the modelling methodology, road degradation models can be divided into three categories, namely the empirical, mechanistic-empirical (M-E) and probabilistic models [35]. The differences among them are explained as following.

Empirical models

A typical empirical model is given in the ASSHTO 1993 design guide [109]. This model adopts a linear combination of factors with coefficients largely derived from the AASHO Road Test conducted in the 1950s and 1960s [181]. Another empirical model is PARIS, a European-wide pavement performance prediction model built upon pavement condition records of fifteen European Union (EU) nations [74]. These models provide empirical relationships between the dependent variables (pavement conditions) and the explanatory

Table 2.3 Pavement degradation models surveyed: full names and abbreviations

Pavement degradation model	Abbreviation	Reference
American Association of State Highway and Transportation Officials (AASHTO) Guide for Design of Pavement Structures	AASHTO 1993	[109, 181]
Performance Analysis of Road InfraStructures model	PARIS	[74]
Highway Development and Management 4	HDM-4	[124]
National Cooperative Highway Research Program (NCHRP) Mechanistic-Empirical Pavement Design Guide	MEPDG	[8, 183]
Whole Life Performance Prediction Model/Long Term Pavement Performance Model	WLPPM/LTPPM	[53]
Arizona Department of Transportation (ADOT) Pavement Management System Model	ADOT model	[94]
Highway Investment Programming System	HIPS	[35, 200]
Highway Maintenance Efficiency Programme	HMEP	[104]

variables (traffic, age, etc.), but the functional forms and coefficients are not derived from any mechanistic principle.

M-E models

M-E models predict pavement conditions by following logical procedures based on the mechanistic principles. Only certain calibration factors are evaluated from empirical data to achieve a better match with real observations. One example of the M-E models is the NCHRP's MEPDG model [8, 183], the current pavement design guide used by AASHTO since 2004. In the MEPDG model, the pavement design life is divided into analysis periods, within which site conditions, including traffic, climate and material properties, are held constant. Based on the site condition inputs, pavement responses (stresses or strains) in each analysis period are then calculated. Next, critical stresses/strains are used to predict distress increments (e.g. monthly increment in rutting depth). Finally, the distress increments are accumulated to obtain the total damages for performance evaluation. This calculation procedure is illustrated in Figure 2.3.

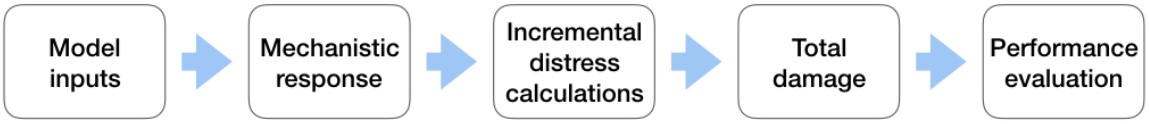


Fig. 2.3 Calculation procedure for the M-E pavement degradation model in the AASHTO 2004 Pavement Design Guide.

The WLPPM developed in the University of Cambridge also incorporates mechanistic theories, but the emphasis is given to the modelling of the dynamic vehicle-pavement interactions and no evidence can be found regarding its application [53].

The widely adopted HDM-4 model stands in the middle of the empirical and M-E models by adopting a structured-empirical methodology, where the functional forms and primary inputs of the degradation equations are decided according to both mechanistic theories and empirical evidence. Statistical techniques are then used to quantify their impacts [124, 158, 170].

M-E models have higher requirements on data quality and also involve more computation work than the pure empirical models. For example, in MEPDG, pavement responses such as the stresses and strains have to be calculated using multi-layer elasticity or the Finite Element method. This is acceptable for the design of several road sections, but becomes cumbersome for network-level analysis.

Probabilistic models

Affected by many factors that are difficult to predict, quantify or even identify, pavement condition prediction is inherently associated with uncertainties. Probabilistic models can address this issue by modelling pavement condition change as a stochastic process, usually a Markov state transition process [57]. An early model embracing such an idea is from the ADOT [94, 188]. In this system, the highway network in Arizona is divided into thousands of one-mile sections and pavement condition change is simulated by multiplying the initial condition vector with the Markov Transition Probability Matrices, or TPMs (T in Equation 2.8). This model influences later practices in places such as the Kansas State, Saudi Arabia and the Nordic countries [200]. Recently, HMEP, a Markov-based pavement deterioration model, is developed and quickly becomes popular among UK local authorities. The toolkit provides default definitions for homogeneous asset groups, condition bands and TPMs for the UK situations but also allows user inputs to override the default values [104].

$$S_n = S_0 \cdot T^n \tag{2.8}$$

Where S_0 and S_n are pavement conditions in Year 0 or Year n expressed as a vector. T is the TPM for Markov-type pavement degradation models.

2.3.2 Inputs used in pavement degradation models

Model inputs reflect model developers' understanding about the influencing factors of the degradation process. Although the definitions, complexity and details of the inputs vary a lot, most models still share similarities. A summary of model inputs and specific measures for each input category is given in Table 2.4.

The basic types of inputs incorporated in almost all models are the time and traffic. Time is usually measured by pavement age or time since last maintenance. Traffic can be expressed by the annual or cumulative Equivalent Single Axle Load (ESAL, a measure of damage to the pavements caused by traffic) or AADT (a measure of traffic quantity). M-E models require the calculation of traffic load spectra rather than a general measure of traffic. Some models use more than one measure for the same concept, such as the four age measures used in HDM-4. However, the ASSHTO 1993 model does not even consider time and degradation is predicted solely as a result of traffic. One particular model, WLPPM, considers the dynamic vehicle effect caused by tyre-surface interaction, but this measure of traffic influence is overly detailed for network-level analysis.

Table 2.4 Inputs of road degradation models

Input	**Measure**	**Used by Model**
Time	Age since construction	Paris HDM-4
	Age since last maintenance	HDM-4
	Age from the start of the analysis	ADOT model HIPS HMEP
Traffic	AADT	HDM-4
	ESAL	AASHTO 1993 PARIS HDM-4
	Others (load spectra, speed and vehicle dynamics)	MEPDG WLPPM
Pavement characteristics	(Adjusted) Structure Number	AASHTO 1993 HDM-4
	Subgrade resilient modulus	AASHTO 1993 MEPDG
	Material properties by later	MEPDG WLPPM
	Others (surface profile and properties, drainage and distress potential)	WLPPM PARIS MEPDG

Table 2.4 Inputs of road degradation models (cont'd)

Climate factors	Temperature	HDM-4 WLPPM HMEP
	Precipitation	MEPDG HMEP
	Moisture/humidity	HDM-4 MEPDG
	Others (wind speed, sunshine, water table)	MEPDG
Maintenance	Reduction of degradation severity	HDM-4 HIPS HMEP
	Reduction of potential defect occurrence	ADOT
Others	Initial or construction quality, initial IRI and month of construction/opening	HDM-4 MEPDG
	Current pavement conditions	AASHTO 1993 ADOT HIPS HMEP
	Road hierarchy	HMEP

The second group of inputs is related to the pavement characteristics. A typical measure is the Structure Number, which expresses the structural strength of pavements [182]. M-E models require complete information on the thickness and material properties of each layer [53]. Markov models usually do not include these measures explicitly, but it is possible to divide "homogeneous groups" based on pavement characters [57].

The third group of input variables describes the climate or environmental conditions that the pavements are subjected to. Two typical inputs are the temperature and precipitation levels. Again, some models may include multiple variables for the same concept, such as average temperature in the coldest as well as the warmest months [53]. HDM-4 uses the climate information to categorise pavements, while M-E models like MEPDG and WLPPM incorporate these values into the calculations.

Another important aspect to consider when modelling pavement degradation is the maintenance effect. Maintenance activities are carried out periodically to correct current defects and to prevent further degradations. Usually in models, the maintenance effects are expressed as reductions on defect severity, and in one case, the reduction on potential for future defects [94]. The former can be easily implemented in Markov-type models by defining maintenance effect TPMs. HIPS model proposes to define the Markov condition bands based on the remaining life before the next maintenance, but no information has been found about the details [200].

There are various other factors that have been studied in pavement degradation modelling, such the construction quality or months since the start of construction and opening to the traffic [8]. All Markov-based models require information about current road conditions to initiate the Markov transition process [57, 94].

2.3.3 Outputs of the pavement degradation models

The outputs of pavement degradation models are the various indicators of pavement conditions. Table 2.5 presents the model outputs and specific measures that are adopted by each model. Some models give results on various defects (e.g., HDM-4 and ADOT), while others only predict the change of a compound index (e.g., HMEP). There is no absolute agreement among different models in regards to the measures of defects. For example, for cracking, a model can either evaluate the depth, length, area, or from other aspects such as the time or traffic load before crack initiation. Despite the obvious increase in complexity, breaking-down of condition prediction into individual types of defects has several advantages, such as identifying major defects that lead to performance losses, selecting suitable maintenance methods to treat corresponding defects and calculating user costs more accurately [124]. One

particular output variable, roughness, is commonly treated as the major influencing factor of road user costs [231].

Table 2.5 Outputs of road degradation models

Output	Measure	Model
Rutting	Depth	Paris HDM-4 MEPDG WLPPM HIPS
Cracking	Initiation time	HDM-4
	Area (in percentage)	HDM-4 HDM-4 MEPDG ADOT HIPS
	Length	MEPDG
	Depth	MEPDG
	Cumulative traffic at the initiation of cracking	PARIS
	Cracking index: weighted sum of length of different types of cracks	PARIS
	Ratio between current load cycles and limit load cycles	WLPPM
Ravelling	Initiation time	HDM-4
	Percentage area	HDM-4
	Ravelling length per unit length	PARIS
Potholing	Initiation time	HDM-4
	Number per unit length	HDM-4
Edge breaking	Volume loss of material per unit length	HDM-4

Table 2.5 Outputs of road degradation models (cont'd)

Roughness	International Roughness Index, IRI	PARIS HDM-4 MEPDG HIPS
	Surface displacement profile	WLPPM
Texture	Depth	HPM-4
	Skid resistance	HDM-4
Compound index	Present Serviceability Index, PSI	AASHTO 1993
	Carriageway Condition Index, CCI	HMEP
User effects		HDM-4 HIPS
Environmental effects		HDM-4
Economic results		HDM-4 MEPDG HMEP

2.3.4 Applicability criteria

Pavement degradation models are usually developed for specific applications. Some are designed to be used anywhere after calibration (such as the HDM-4 model), while others are only suitable for specific regions (such as the PARIS model which is only applicable within the EU nations). Quite a few are designed for the UK, such as the HMEP model. HDM-4, developed by the World Back, is probably the most popular pavement condition prediction model worldwide and has been used not only in the asset management stage, but also for road project appraisal. For models designed to be widely applicable, there is still a need to calibrate the model based on historical observations in the application area before they can be utilised by highway engineers and asset managers with confidence [7].

The deployment time periods of different models also vary. WLPPM, which was developed in the 1990s, seems to have never come into application. While the AASHTO 1993 model was replaced by the MEPDG model in 2004. HDM-4, developed in the late 1990s, is still widely used today. For models that are outdated, their technical details (e.g., modelling methods, functional forms) may still be valid. An example is the Markov chain

based ADOT model. Although it has been replaced by more advanced models [239], its major contribution, namely modelling the degradation process as a Markov chain, is still widely adopted in newly developed models, including the HMEP model developed in the UK after 2000 [104]. The problem of outdated models is usually the data. For instance, the ASSHTO 1993 model was based on road test experiments in the 1950s and 1960s [181]. These experiments may not represent the current situations of traffic load or pavement design. This illustrates that when developing new models, it is acceptable to adopt the modelling methods and functional frameworks from existing models, although specific coefficients and values need to be adjusted to reflect the present conditions. Given the diverse ranges of applicability of the existing models, the validity of the adopted model should be evaluated before the implementation.

Table 2.6 Summary of road degradation models

Model	Type[1]	Region	Time	Inputs						Outputs										
				Age	Traffic	Pavement	Climate	M&R	Others	Rut	Crack	Ravel	Pothole	Edge	Roughness	Texture	Comp.[2]	User	Environment	Economic
AASHTO 1993	EMP	USA	1993		●	●			current condition								PCI			
PARIS	EMP	EU	1998	●	●					●	●	●			●					
HDM-4	STR-EMP	World	2000	●	●	●	●	●	construction quality	●	●	●	●	●	●	●	●	●	●	●
MEPDG	M-E	USA	2004		●	●	●		initial quality	●	●				●					●
WLPPM /LTPPM	M-E	USA	1995		●	●	●			●	●				●					
ADOT	PROB	AZ, USA	1980	●				●	current condition						●			●		
HIPS	PROB	Nordic	1990	●				●	current condition	●	●				●			●		
HMEP, NETCOM, STRAT-2	PROB	UK	2012	●			●	●	current condition, road hierarchy								CCI			●

[1] EMP: empirical model; STR-EMP: structured-empirical model; M-E: mechanistic-empirical model; PROB: probabilistic (Markov) model.
[2] Composite pavement condition index. PCI: Present Serviceability Index; CCI: Carriageway Condition Index.

2.4 Summary

The literature review begins with a brief review of the sustainable transportation concepts in Section 2.1. Under the sustainability assessment framework, the performance of a transportation system is evaluated by various indicators belonging to the social, environmental and economic aspects. Although the focus of this thesis is not on carrying out a comprehensive sustainability analysis, the model-based indicator quantification and system evaluation methods from the sustainable transportation literatures are still found to be beneficial and worth referencing for the analyses in this study.

In order to forecast the traffic states of a transportation network under future scenarios, various traffic simulation models are reviewed in Section 2.2. Specifically, traffic dynamics are considered to be generated through the interactions of the travel demand and the network supply. Travel demand modelling approaches include, from simple to complex, the trip-based, tour-based and activity-based methods, where the spatial and temporal units become increasingly disaggregated and the analysis period more extended. In terms of the network supply, it can be modelled according to the macroscopic, mesoscopic or microscopic traffic propagation theories. Macroscopic models have the advantage of being efficient by considering aggregated traffic characteristics. Microscopic models, on the other hand, have high fidelity as the movements of individual vehicles are modelled at a cost of the computational speed. The mesoscopic models combine the features of the macroscopic and microscopic approaches. In mesoscopic traffic propagation models, individual vehicle routes and choices are retained, but the link-level dynamics are simplified into aggregated quantities.

In Section 2.3, existing pavement degradation models are reviewed. It is found that they can be categorised into statistical, M-E or probabilistic models depending on the nature of the underlying degradation relationships, i.e., whether the mechanistic principles and stochasticities are considered. Pavement age and traffic load are found to be the most common model inputs, while other inputs include the design and material characteristics, the climate, the maintenance and other factors. In terms of the outputs, most pavement degradation models lead to estimations of the severity of different types of defects (e.g., rutting, cracking, roughness and potholing). Two of the models reviewed also output composite indices of pavement conditions. In addition, each pavement degradation model has its specific applicable area, which should be considered before being adapted to other cases.

The review of the sustainability indicators, the traffic and pavement condition prediction models will assist the development of the methodologies in the rest of this thesis. Currently, the traffic and pavement condition models in the literatures are not well integrated. In order to combine these two types of models, priority should be given to the technical compatibility between them (e.g., the connections between the modelled variables, the spatial and temporal

resolutions). In addition, it is necessary to ensure that the adopted models can be computed efficiently to calculate city-scale network-wide indicators/metrics. As a result, simple but efficient models could be more favourable to this end. Given these general thoughts, the detailed reasoning behind model selections will be discussed in corresponding later chapters, after explaining the practical constraints imposed by data availability in the case study area and the computational expenses.

Chapter 3

Mesoscopic Model for City-Scale Traffic Simulations

3.1 Introduction

In this chapter, a mesoscopic traffic simulation model with trip-based disaggregated travel demand will be developed. The aim of developing this modelling tool is to conduct efficient traffic simulations on the detailed city-scale road network of the study area. The model inputs are based on several open data sources, including the OpenStreetMap (OSM), the United States Geological Survey (USGS), research findings on the travel demand of Transportation Network Companies (TNC, including Uber and Lyft) released by the San Francisco County Transportation Authority (SFCTA) as well as other open data channels. The mesoscopic traffic model is designed to run on the High Performance Computing (HPC) clusters, thereby improving the computational speed significantly. Besides, a preliminary validation is conducted by comparing the simulation outcomes with the published results by the Bay Area Metropolitan Transportation Commission (MTC) traffic model.

This chapter is organised into the following sections. First, the background will be introduced through examples of previous research targeted at the same study area, including PhD research as well as existing models developed by the local transportation agencies. Features, pros and cons of each model will be highlighted. Although the existing models have been very useful for their respective purposes, they often do not have the suitable level of efficiency or spatio-temporal resolution required by this research. Consequently, a new mesoscopic traffic simulation model framework is proposed, which retains the highly detailed network, disaggregated travel demand but with simplified traffic dynamics and traveller behaviours. In section 3.4, the input data to build the traffic model from open

data sources will be introduced. This will be followed by the introduction about the model structure in Section 3.5. In the end, results from the simulation will be presented, followed by the results of a preliminary validation of the model outcomes based on a published model in Section 3.6.

The adopted traffic model uses disaggregated vehicle-level travel demand, an important feature that facilitates the inclusion of vehicle-level choices and behaviours (e.g., route choice criteria). While this also implies that the model to be developed needs to find the optimum path for each of the millions of vehicles in the simulation. Exact path finding algorithms such as the Dijkstra's algorithm [67] are generally computationally intensive. As will be introduced in Section 3.5, route calculations for vehicles that belong to the same route assignment sub-step can be carried out in parallel. This provides an opportunity of accelerating the simulation by adopting parallel computing. The HPC facility is adopted for this purpose. There are 32 cores on each computation node of the adopted HPC facility, thus the route finding algorithm can spawn 32 parallel processes simultaneously. To avoid the data communication overhead of running on multiple nodes, the current programme is limited to only one node of the HPC.

3.2 Background

The case study area of San Francisco (SF) has been the test bed for many transportation related innovations. In the field of traffic simulation, it has long been incubating pioneering models and applications. The city is facing daily urban congestion issues as well as long-term threats of the earthquake hazards, where simulation tools are helpful for planning and testing various scenarios. In addition, some model implementations are particularly targeting at emerging technologies, such as the electric vehicles. This section surveys some of the most recent work developed for the city of SF and the whole Bay Area, with the objective of understanding the methodologies, the targeted applications of each model as well as to cumulate knowledge of potential data sources for the improvements or developments of new models.

Bay Area User Equilibrium (UE) model by Sheehan [217]

The model developed by Sheehan [217] is with a clear goal of finding the network flow that follows the static UE principle, where no travellers can decrease his or her travel time by following a new path [60, 244]. Road link characteristics, including the numbers of lanes, speed limit and capacity are obtained based on OSM default values. Link travel time follows the macroscopic volume delay relationship from the Bureau of Public Roads (BPR). The UE

solver is based on Steele [222]. The framework of this traffic model also provides functions for alternating the network, such as decreasing the capacity, to simulate the effect of flooding or other disruptive events. Figure 3.1 shows the example output of the model for the Bay Area road network (containing secondary roads and above). The static assignment procedure is similar to the one that is ultimately adopted in this study. However, as the interest of Sheehan [217] is to model the equilibrium state traffic distributions on the highways and major roads, the travel demand profile is also much simpler: the travel demand in Sheehan [217] is aggregated to Traffic Analysis Zones and a single time slice, rather than the hourly demand profile precise to each road intersections used in this study.

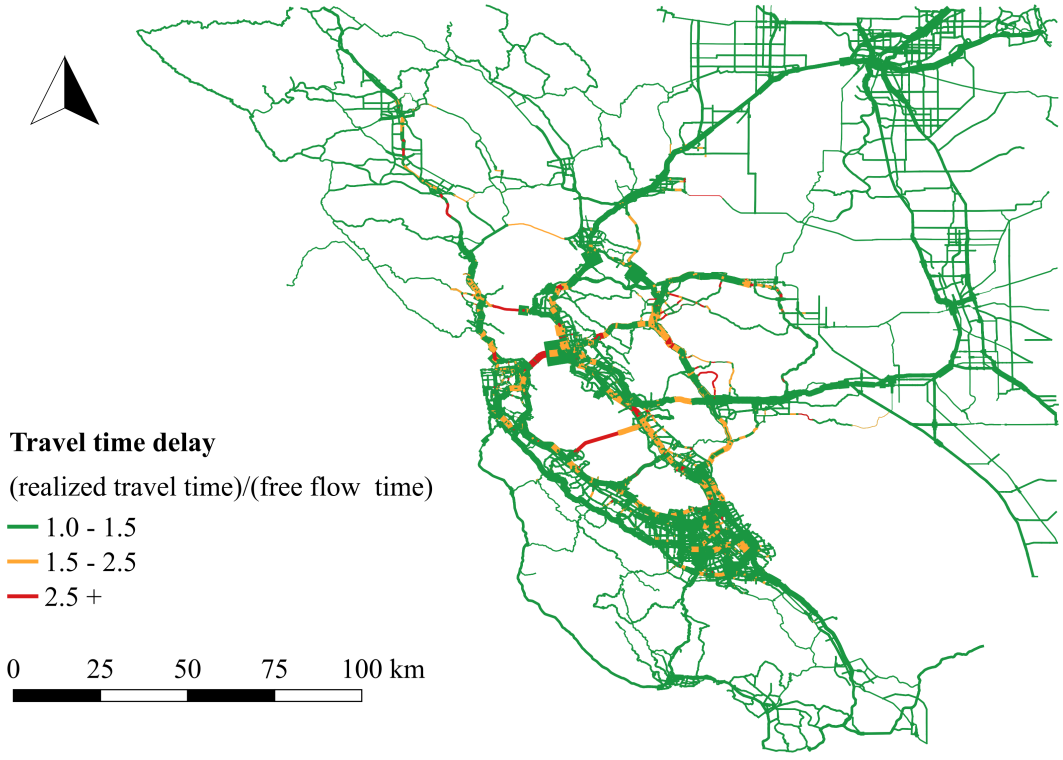


Fig. 3.1 UE traffic flow by Sheehan [217]

Bay Area MTC Travel Model One

Travel Model One is the activity-based traffic model developed for the Bay Area MTC. The model emphasises more on the demand modelling side, which allows the users to produce realistic activity patterns and travel behaviours of individual decision makers. Individual travel decisions are made according to the socio-economic status of each individual/household

and other factors such as the coordination between household members [73]. The travel demand model is paired with a multi-class static user equilibrium solver to obtain the traffic volumes for five periods of a day (early morning, morning peak, mid-day, early evening and evening peak). In this sense, although the travel demand is based on one-hour time slice, the temporal resolutions of the final results are limited to the five periods used in the assignment model [160]. Compared to the trip-based travel demand used in this thesis, Erhardt et al. [73] adopted a more advanced activity-based demand generation, which could be an area of further improvement.

Bay Area model by Miller [154]

In the PhD thesis of Miller [154], the MTC network and demand for the Bay Area are used to evaluate the transportation system performance under seismic damages. The road network consists of around 10,000 nodes and 25,000 edges for the whole Bay Area (in this chapter, a graph of similar size is used for representing the SF network, a city inside the Bay Area). An efficient traffic model is coded from scratch for traffic assignment. The average hourly travel demand in the morning peak is used in her case study. However, the origins and destinations of the travel demand is greatly simplified to 34 supernodes distributed in the whole study area. To model more complex scenarios, a second high fidelity model, the MTC's activity-based model, is adopted. The high fidelity model includes not only transit networks made of local bus, ferry and rail routes, but also the variable travel demand (change of destination or forgoing a trip) occurring after an earthquake. However, because of the detailed activity based demand generation and the iterative-based UE solver, it takes several hours to run the high-fidelity model to evaluate the traffic conditions under each scenario. The time-consuming high-fidelity model is clearly not a good choice for the city-scale traffic scenario analyses in this study (due to the time constraints of simulating many scenarios). While the spatial resolution in the efficient model (34 super nodes) is also too simplistic for modelling individual routing preference in this study.

BEAM

BEAM (Behaviour, Energy, Autonomy and Mobility) is a collection of software tools developed at Lawrence Berkeley National Laboratory (LBNL) for the simulation of transportation and electric system, such as the charging infrastructures and behaviours of plug-in electric vehicles (PEV) [218]. The traffic simulation in BEAM is largely based on the multi-agent simulation tool MATSim. MATSim models each driver explicitly as one agent. Traffic flows generated by these agents are obtained from queue-based simulations, where agents

go through each link at free flow speed or have to queue if it becomes congested [13, 14]. Once the UE solution has been found by MATSim, BEAM allocates extra electric vehicle drivers to the network to finish their predetermined itineraries while evaluating the necessity and choosing locations of charging their PEVs. The BEAM framework has been applied to a down-sampled population of 463,000 agents in the Bay Area, among which 59,000 are PEV owners. PEV charging demand profile PEV has been compared against observed data [218]. The agent-based approach adopted by BEAM/MATSim is a good direction for this research in terms of moving towards more dynamic traffic analysis.

SFCTA toolbox

The traffic planning models used by the SFCTA are among the first activity-based models extensively used in practice. It was created by traffic consulting firms Cambridge Systematics and Parson Brinckerhoff (now WSP USA) for the SFCTA in 2000 [41]. It features sophisticated demand generation functions that provides detailed full-day activity planning based on a synthetic population of the city, including destination choice, travel time choice and mode choice. Matching the detailed demand generation is a suite of assignment tools, including a static user equilibrium tool (regional-level), a dynamic traffic assignment tool (city-level) and a highly realistic microsimulation tool (project/corridor-level). One particularly interesting aspect in the SFCTA traffic model toolbox is the city-level dynamic traffic assignment tool (DTA), which can simulate time-dependent traffic flow for every road in the city and produce more realistic results than static traffic assignment (STA, where there is no notion time). However, as DTA inherently relies on iterations, for a five hour scenario simulation with 620,000 vehicles, the model converges after 20 iterations and takes around 50 hours for computing [202]. The SFCTA model is highly specialised and has many advanced features, however, the speed of the simulation may limit its application for analysing large numbers of scenarios, and the complexity may hinder its adoption and increase the difficulty of adaptation for interdisciplinary research projects.

As a summary of the city-scale simulations and analyses reviewed above, it is desirable to include a detailed representation of the road network and a small enough temporal resolution to reflect the dynamics of the busy transport network. However, on the other hand, it is also crucial to curtail the computational time when there are hundreds of scenarios to be analysed such as in the traffic and CO_2 emission analysis presented towards the end of this thesis. To resolve this contradiction in scale, resolution and computational effort, a new mesoscopic traffic model is developed in this chapter. It features a full representation of the city's drivable road network, hourly varying travel demand and a sub-step incremental assignment (an intermediate solution between STA and DTA) that can run in parallel. Detailed description

of the model structure, inputs and outputs will be presented in this chapter, together with a preliminary validation of the results.

3.3 Motivation

In this chapter, a traffic simulation model is developed with the primary aim of balancing the efficiency and fidelity that are both necessary for city-scale simulations. From the literature review of traffic simulation models, it can be seen that many advanced models and theories have been developed based on some long-established simpler representations of the problem. For example, from the trip-based demand modelling to activity-based models. In this thesis, however, it is decided that the simpler theories of traffic modelling will be used, as long as it is sufficient to address the research question. Below are a few justifications of this decision. First of all, the classical models (e.g., STA, trip-based demand models) are usually more computationally efficient, a feature desired for analysing large-scale problems (e.g., city-scale simulations). Secondly, as the traffic model is going to be combined with the pavement infrastructure model for the transportation system performance simulations, it is more important to ensure that the two interdisciplinary models have matching scales and resolutions, rather than one being much more sophisticated than the other. For example, both of the traffic simulation model and pavement degradation model used in this study consider road links as their basic spatial units. More detailed sub-link vehicle dynamics or metre-by-metre pavement degradations are not included. Thirdly, starting from the classical theories, it facilitates the understanding of the fundamentals of the problem while improvements (e.g., sub-link vehicle dynamics) can be carried out in the future work.

To be more specific, in terms of the spatial complexity, a detailed topology of the street network up to the link-level is retained, but not the sub-link features, such as individual lanes. This is because to model sub-link behaviours such as lane changing, time-consuming microscopic simulations are needed, which will inevitably make the model less efficient for city-scale analysis. While link-level models are commonly adopted for various types of traffic simulations, including emission calculations, as shown in Zeng et al. [252].

Secondly, for the representation of the travel demand, it should be disaggregated into individual trips in the study area, so as to allow certain individual behaviours (e.g., route choices and eco-routing behaviours) to be modelled. As an initial attempt for simulating city-scale traffic patterns, individual information other than the origin, destination and departure time information is not included. This level of abstraction is sufficient for modelling certain traffic operation and eco-friendly travel scenarios, although it is acknowledged that it is

Table 3.1 Features of the traffic model after weighing the efficiency and fidelity in traffic simulation

Features	Included	Not included
Network representation	Topology and link-level attributes of all drivable roads, including the freeways, arterials and down to the residential streets	Curves and bends, short bottlenecks, individual lanes, etc.
Travel demand	Individual trip's origin, destination, departure time and route choice criteria	More detailed personality or socio-demographic background of the individuals
Temporal dynamics	Static assignment based on one-hour time step and many sub-steps	Second or minute level traffic distributions based on DTA

desirable to include trip time choices or even individual activity patterns in the future for greater modelling flexibility (at a cost of the computational complexity).

Lastly, in terms of the temporal discretisation, static assignment with small assignment period will be used, again due to the cost of truly dynamic assignment models. Static assignment is appropriate, as the simulation of the city-scale CO_2 scenarios do not always require the understanding of minute-by-minute or second-by-second traffic patterns. A suitable time slice size should be short enough to capture the temporal evolution of the traffic patters, while also long enough to make the simulation computationally feasible. Traditionally, transportation planning models adopt a time interval of several hours (e.g., 3-hour long for the morning peak period, 8 hours for the off-peak period). However, given that the size of SF is not very big (11 km by 11 km), a one-hour time slice is also feasible, as most journeys can finish in one hour even during the most congested time of the day. For time intervals shorter than one hour, it will break the assumption of STA that some trips cannot be finished in one time step. To stabilize the traffic assignment within a time step, sub-steps are also involved when trips are assigned to the network incrementally and the network conditions (e.g., congestion level) are updated at the end of each sub-step. In this sense, the final flow distribution will be an approximation of the UE, where the true UE refers to the situations that no user can find a shorter/faster travel time by unilaterally changing his/her route [60, 244]. Convergence to dynamic user equilibrium is not enforced due to (1) there is no guarantee that dynamic user equilibrium does exist in real life; (2) the convergence to user equilibrium requires many iterations and is very often the most time consuming part of traffic simulations.

The model will be referred to as mesoscopic because it explicitly models individual trips in the network but not their sub-link behaviours. In this chapter, the route choice behaviours

of different travellers are assumed to be uniform, i.e., all travellers will choose to use the time step-dependent shortest path. This assumption will be relaxed towards the end of the thesis, when the carbon emissions will be included into the route choice decisions of some travellers. A summary of the features of the traffic model is given in Table 3.1.

3.4 Model inputs

Two types of data are required for the mesoscopic traffic simulations: the network properties (topology, capacity, speed limit, ...) and the travel demand (origin, destination, departure time, ...). This section details the process of data collection from openly available data sources as well as the procedures involved to clean the data.

3.4.1 Network supply from the OSM

OSM is a popular service that offers free editable digital map of the world. Map data on the OSM come from public domain mapsets, licensed aerial imageries or GPS tracelogs uploaded by volunteers [172]. In the US, where the study area is located, the OSM road network was initially populated by the public domain TIGER maps in 2007/08 and has been gradually updated by the community over the years. The good level of completion and standard data format have made the OSM a useful network dataset used in many previous transportation studies, including the microscopic traffic simulation package SUMO [16] and OSMnx, a comprehensive Python tool for downloading, cleaning, analysing and visualising street networks [27].

Downloading the OSM road network

The OSM road network can be conveniently filtered and downloaded with the Overpass API [175]. Figure 3.2 shows the Overpass Query Language (QL) script used to download the road network for SF. Map features in the OSM are denoted by tags, such as “building”, “railway”, “power”, “waterway”, etc. In particular, the “highway” tag identifies all roads and footpaths, which not only includes the real highways but also residential roads and pedestrian-only paths [173]. The “way[highway]” in the QL script selects all the roads in the given bounding box. Figure 3.3 provides an example of the downloaded road network data in the JSON (JavaScript Object Notation) output format.

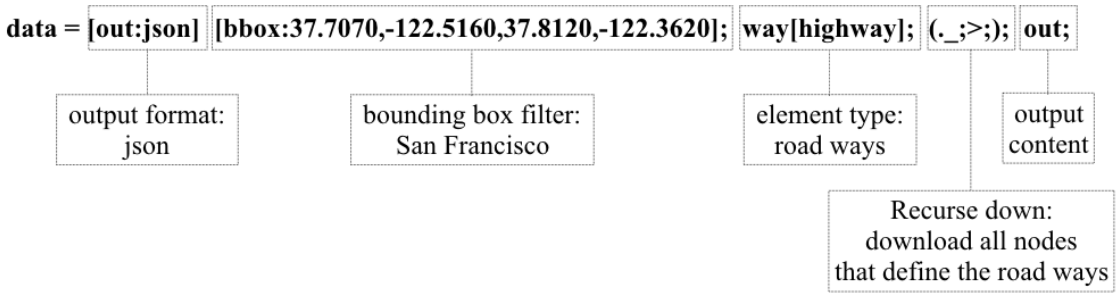


Fig. 3.2 Overpass QL to download the SF road network.

Data cleaning procedures for the OSM road network

Data downloaded from the OSM contain useful information on the topology and attributes of a traffic network, e.g., speed limits and lane counts. However, the raw data also include redundant and/or missing details that should be handled. Overall, the OSM network needs to be converted into a concise, directed graph so that subsequent traffic simulations (e.g., the shortest path finding algorithm) can be run efficiently. In this section, the pipeline for converting the raw OSM data into a directed graph is explained. It consists of 3 steps: (1) removing redundant nodes and constructing a road network graph from the OSM data; (2) adding directionalities; (3) populating graph properties.

A graph made of nodes and links (also called vertices and edges) is a useful abstraction for representing networks in the real life. Examples include a graph representation of the road network or a social network [64]. For road network graph, each graph edge corresponds to an individual street between two intersections (graph vertices). Graph edges can be weighted to reflect street properties, e.g, time to traverse the street. Also, a road network graph is usually directed, i.e., the weight of an edge is infinitely large opposite the normal traffic flow directions. In the raw OSM data, unlike an edge of a graph that connects exactly two nodes, the OSM “ways” string together multiple nodes. An example of this is shown by the list of "nodes" for a "way" element in Figure 3.3. Consider the OSM way element that represents the famous Lombard Street in SF (Figure 3.4(b)). It contains 150 nodes shown as the (red) dots in Figure 3.4(b), but only the first and last nodes are meaningful intersections to adjacent roads. The rest 148 nodes only serve to depict the geometry of the hairpins ("geometric points"). As the speed of many graph computation algorithms strongly depends on the number of nodes and edges in a graph, it is necessary to remove these "geometric points" and obtain a more concise representation of the road network. The 150-node way element in Figure 3.4(b) is eventually replaced by a direct link between the start and end nodes, with the new link inheriting the properties of the original way element. On the other

```
{
 "version": 0.6,
 "generator": "Overpass API 0.7.55.4 3079d8ea",
 "osm3s": {
  "timestamp_osm_base": "2018-09-10T15:44:03Z",
  "copyright": "The data included in this document is from
www.openstreetmap.org. The data is made available under
ODbL."
 },
 "elements": [


{
 "type": "node",
 "id": 26117861,
 "lat": 37.6539819,
 "lon": -122.4071324
},
…
{
 "type": "way",
 "id": 394222581,
 "nodes": [
  704673274,
  3972323851,
  3972323849,
  3972323848,
  3972323847,
  3972323846
 ],
 "tags": {
  "highway": "tertiary",
  "name": "King Drive",
  "oneway": "yes",
  "tiger:cfcc": "A41",
  "tiger:county": "San Mateo, CA",
  "tiger:name_base": "King",
  "tiger:name_type": "Dr"
 }
},
…}
```


Fig. 3.3 Example of downloaded OSM data in JSON format.

hand, if some node in the middle of a way element is an intersection that connects with other roads, the way element is split into multiple links at these locations. To accurately represent the actual travel distance, the length of the new link is calculated based on the original unsimplified "way" element, rather than using the straight line distance between the final nodes. The second type of redundant nodes to be removed from the raw OSM data is the “fake intersection”. These are different from the “geometry points” which are in the middle of one way elements. “Fake intersections” are nodes shared by two way elements. For instance, the Fell Street is represented as two connected way elements in the block shown in Figure 3.4(c). However, travellers starting from A are allowed no other choice at B but to go to C. Thus, the presence of node B is not necessary and is omitted in the road network graph. Instead, a new link A-C is created containing the aggregated properties of the original A-B and B-C. After removing the redundant nodes and fake intersections step, an undirected graph of the SF road network is created. This process has also been done in the excellent package OSMnx [27]. It is chosen to be done from scratch here for greater flexibility and control of the cleaning process for this study.

The directionality of an edge in an OSM road network is encoded using a tag called “oneway” and not merely by the order of the “nodes” forming an edge. A way element with an “oneway” affirmative tag value of “yes”, “true” or “1”, it indicates the first node is indeed the entrance to this road and the last node is the exit. If the value is “reverse” or “-1”, then the directionality of the edge is defined with the order of the associated nodes reversed. When the value of the “oneway” tag is “no” or missing, the edge is considered two-way accessible (default value). In this case, the original edge is replaced by two new edges with the opposite node orders. This process is shown in Figure 3.5 and it allows the creation of a directed graph of the SF road network, with each edge representing a specific direction explicitly.

Road attributes, such as lane counts and speed limits, if missing, are set to the OSM default values [172]. The free flow travel time of an edge is set according to the speed limit, with additional delays at the intersections. It is assumed that the traffic signals on highways are synced by car speeds and will not cause delay. Assumptions and default values for imputing missing information are given in Figure 3.6. Figure 3.7(a) shows a cleaned network for the study area, with 9,643 nodes and 26,893 edges. While Figure 3.7(b) offers a visual comparison of the simplified network with the raw OSM data.

Capacity calculation

The most widely accepted capacity calculation is based on the Highway Capcity Manual 2010 (HCM 2010) [229]. However, it is not possible to obtain all the detailed inputs required from open data sources like the OSM for using the capacity calculation formulae in

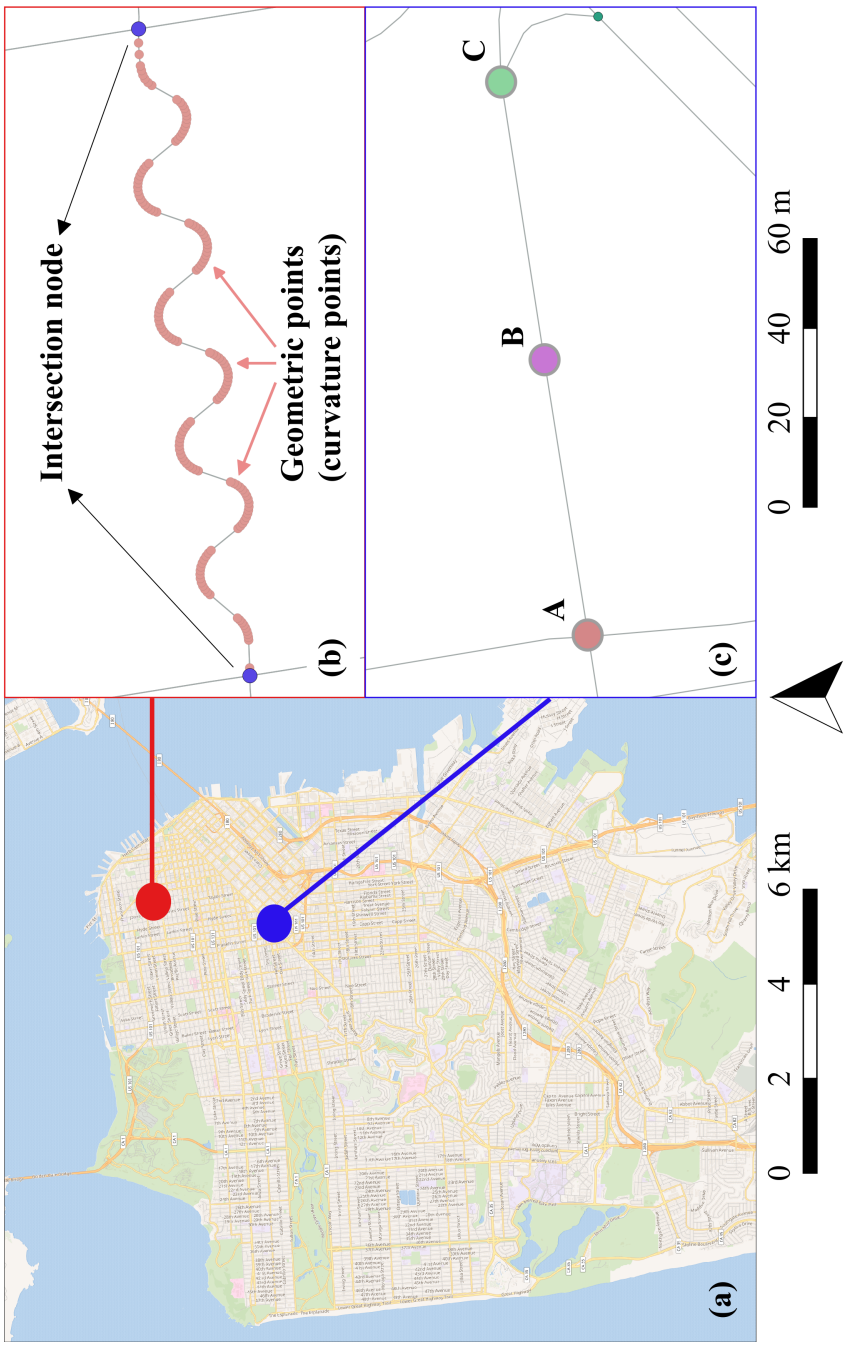


Fig. 3.4 Examples of geometric points and fake intersections to be removed during the data cleaning stage. (a) Map of SF. (b) Intersection nodes and geometric points on Lombard street, where the geometric points should be removed to obtain a simplified road network graph. (c) Fake intersections on Fell street to be removed (Point B).

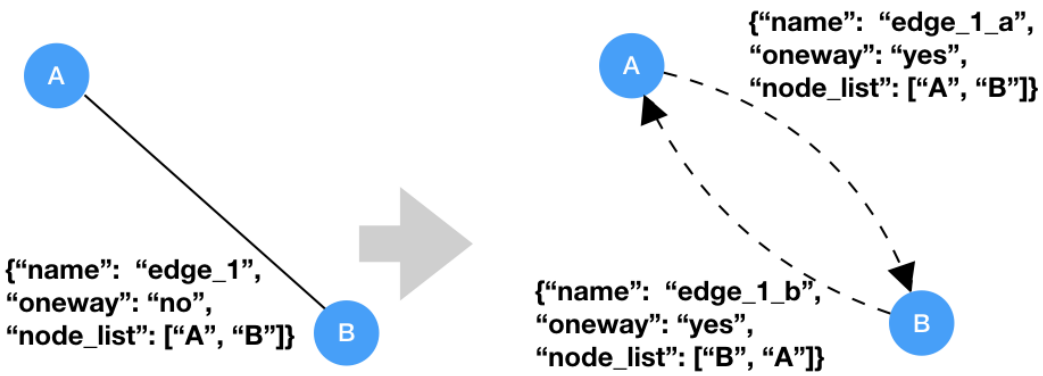


Fig. 3.5 Adding directionality to two-way roads.

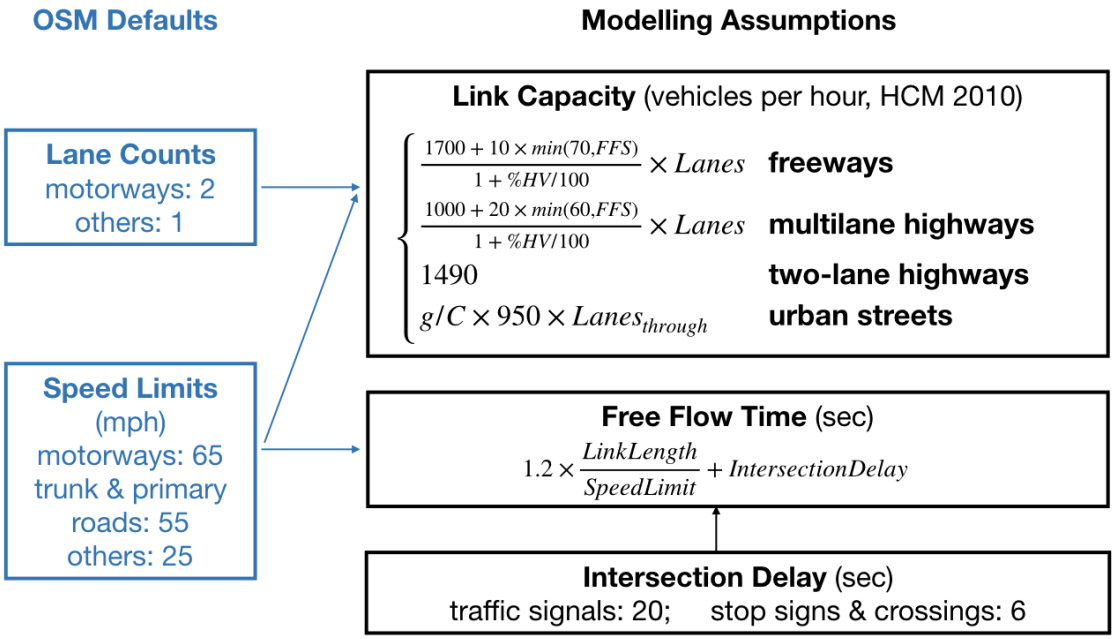


Fig. 3.6 Calculating edge attributes.

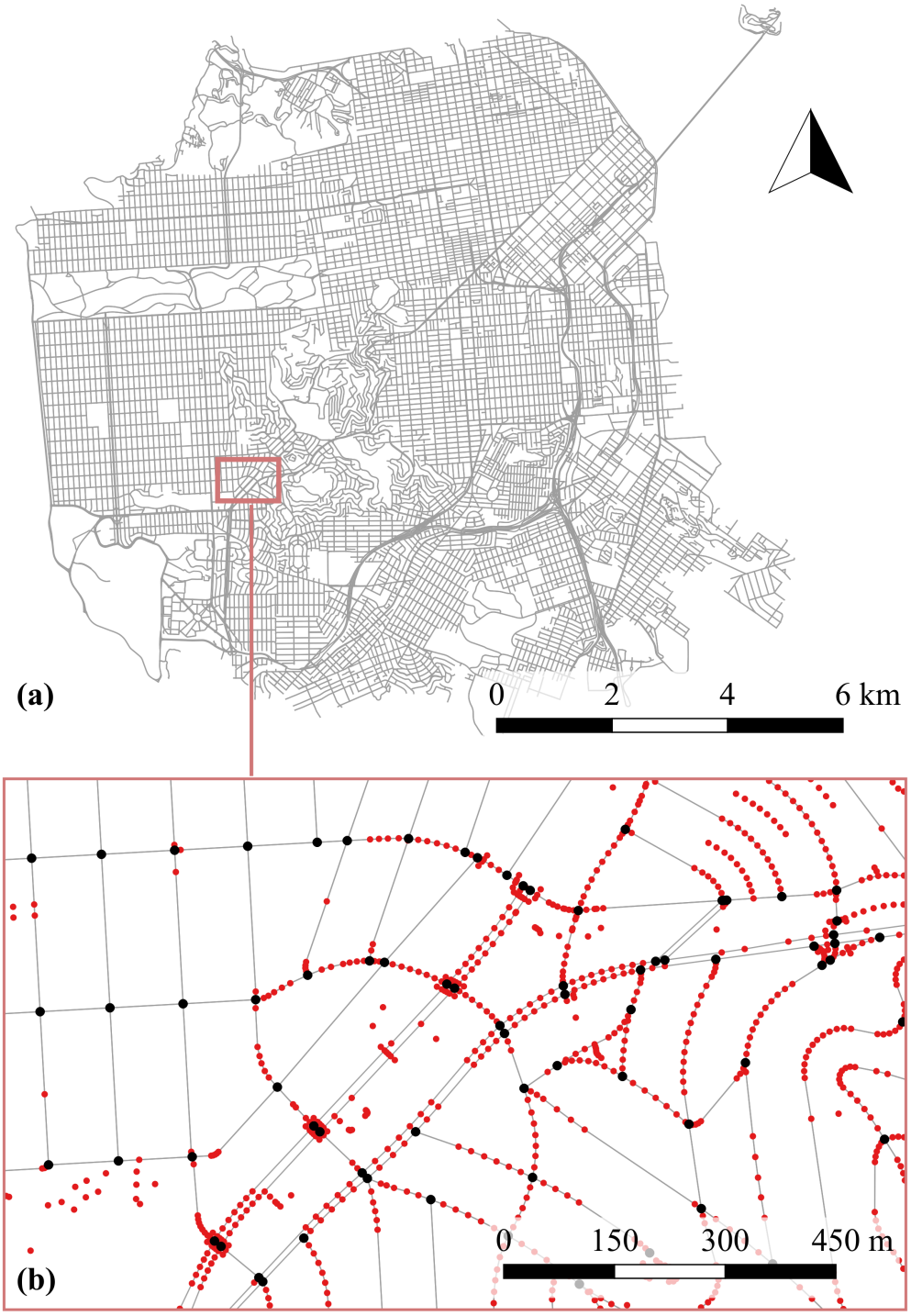


Fig. 3.7 Simplifying the OSM network. (a) Simplified OSM road network for SF. (b) Comparing nodes before and after the network simplification. Red: nodes in raw OSM data; black: nodes in the simplified network

the HCM 2010. National Cooperative Highway Research Program (NCHRP Report 825) provided a simplified capacity calculation procedure based on Highway Capacity Manual 2010 and is adopted here [69]. The translation of OSM road classification to the US highway classification is based on Sheehan [217]. The calculation of capacity (vehicles per hour) is shown as the following.

1. Freeways (motorway and trunk road in OSM)

$$capacity = \frac{2200 + 10 \times (min(70, FFS) - 50)}{1 + \%HV/100} \times Lanes$$

where FFS = Free flow speed (mph), use OSM speed limit

$\%HV$ = Heavy vehicle ratio, use default value for urban roads: 5%

2. Multilane highways (primary road with lane ≥ 2 in OSM)

$$capacity = \frac{1000 + 20 \times (min(60, FFS))}{1 + \%HV/100} \times Lanes$$

where FFS = Free flow speed (mph), use OSM speed limit

$\%HV$ = Heavy vehicle ratio, use default value for suburban roads: 5%

3. Two-lane highways (primary road with lane = 1 in OSM)

$$capacity = 1490$$

4. Urban streets (secondary road and below in OSM)

$$capacity = g/C \times N_{TH} \times s$$

where g/C = effective green ratio of traffic light, use default: 0.45

N_{TH} = number of through lanes

s = saturation flow rate for the through movement, use default = 1900

Adjusting capacity by road gradient

SF is known for its hilly terrain. OSM does not usually carry the road grade feature, but there are some free or paid-for services that provide point elevation, such as the Google

Elevation API, the Open Elevation API or Digital Elevation Map (DEM) from the USGS. In this study, the 1/9 arc-second DEM dataset from USGS is used. The 1/9 arc-second is its highest resolution available from the USGS and is equivalent to a raster cell size of around 3 metres. The two end points of each road are first mapped to this DEM map to obtain the elevations. Based on the elevation difference and the horizontal length of the road, the road gradient can be calculated. Figure 3.8 shows the DEM of the study area as well as the elevation distribution of the road network graph vertices. In addition, the length and free flow travel time of the road are also updated by taking into account the height difference.

The gradient of the road is considered to affect the hourly capacity based on the HCM 2010 [229] and the adjusted capacity after taking gradient into account is given by Equation 3.1. Other researchers have found a stronger influence of gradient on capacity [20], but the HCM 2010 equation is still adopted for its wide acceptability. Figure 3.9 shows the calculated capacity after adjusting for the road gradient.

$$\begin{aligned} &gradient\ \ adjusted\ \ capacity \\ &= unadjusted\ \ capacity \times (1 - \frac{\%gradient}{200}) \end{aligned} \tag{3.1}$$

3.4.2 Intra-city travel demand from aggregated data

The OSM data offers a highly accurate representation of the traffic network. However, travel demand data of matching levels of detail are hard to obtain. One possibility is to generate a synthetic population based on the demographic background as in activity-based models [84, 174]; another method is to extrapolate zonal-level travel surveys to obtain node-to-node travel demand. The latter approach is adopted in this study to take advantage of existing datasets.

Researchers at the SFCTA and the Northeastern University collected, analysed and released data on passenger pick-ups and drop-offs by the TNCs in SF [213]. Through their analysis, they reported that the TNCs account for about 15% of all intra-SF vehicle trips, making it moderately representative of the overall traffic pattern. Despite the rather short data collection period, this dataset is utilised to inform the intra-city travel demand due to its representativeness and high spatio-temporal resolutions (Table 3.2). One drawback of the data is that it only includes trips that start or end in SF, namely the intracity traffic. This limitation can be overcome to some degree by considering the traffic that enters and exits SF through four major entry points, as will be explained in the next subsection.

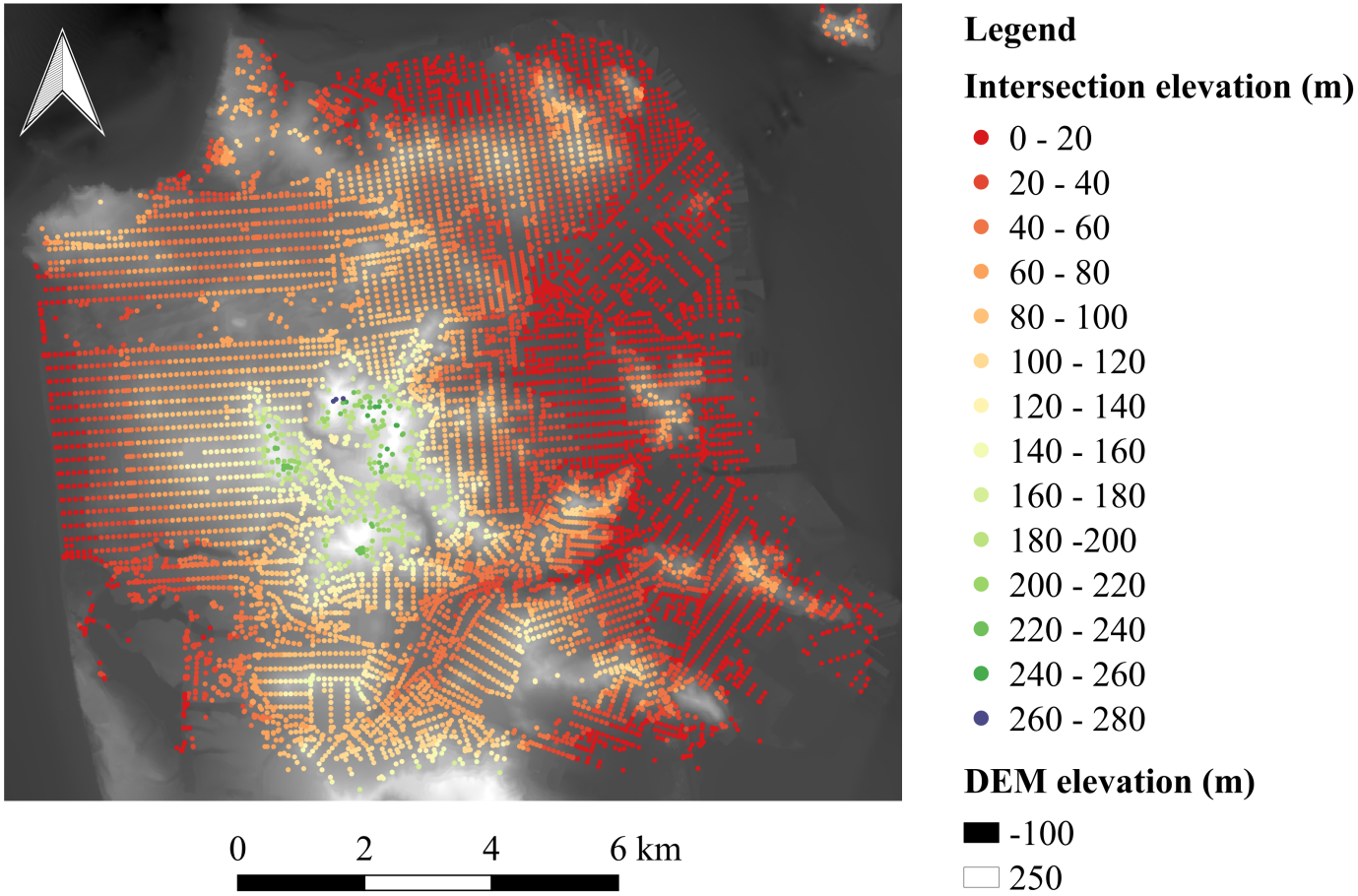


Fig. 3.8 Elevation of SF and its intersections

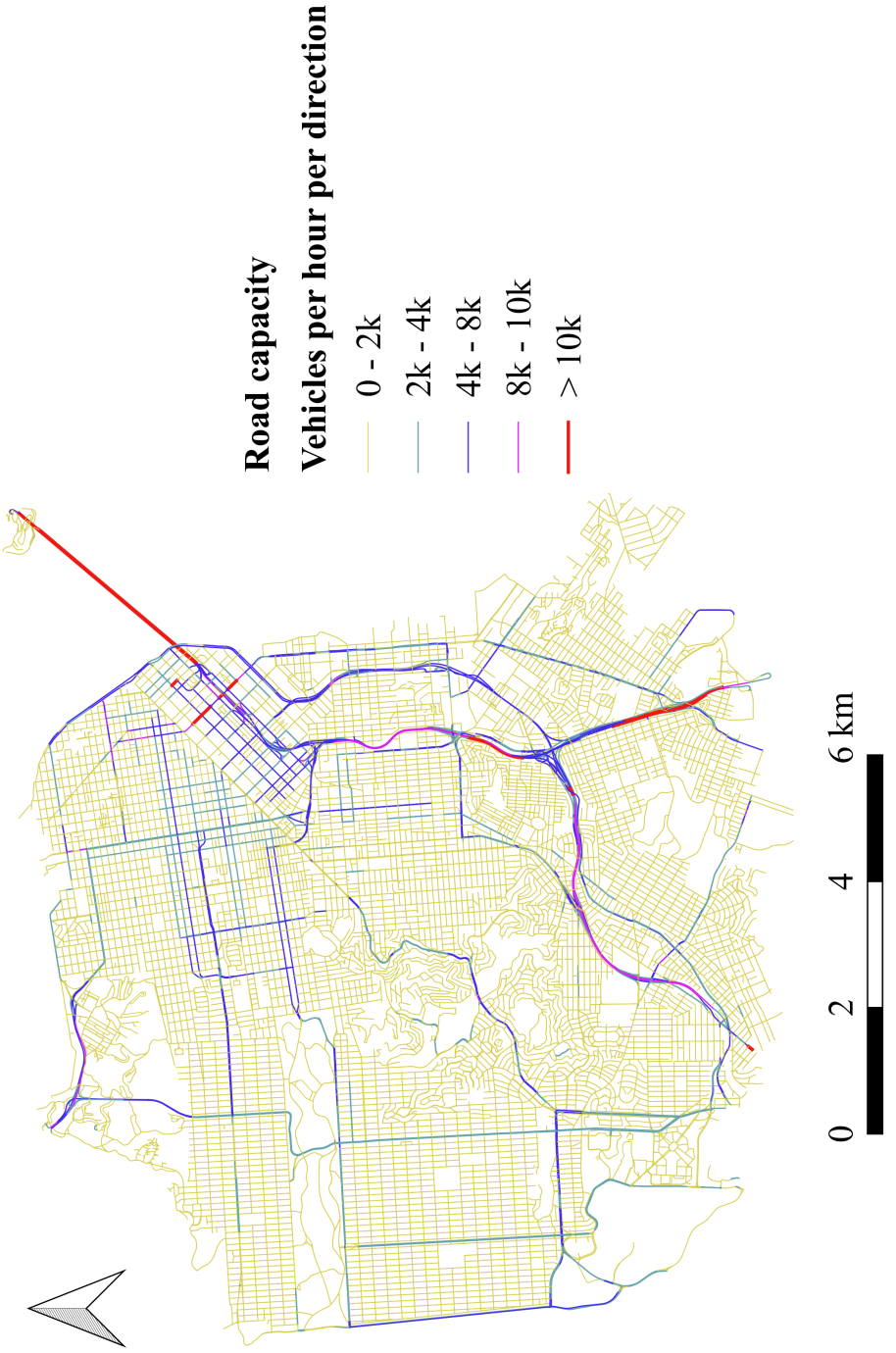


Fig. 3.9 Map of road capacity distribution in SF

Table 3.2 Pros and cons of SFCTA TNCs data

	Spatially	Temporally
Pros: high resolution	Origins & destinations associated with 981 TAZs, each containing 4-8 blocks	Origins & destinations associated with a specific hour of the week.
Cons: limited coverage	Only trip starts & ends in SF are considered.	Data collected in 6 weeks: mid-Nov to mid-Dec, 2016.

Generating disaggregated nodal-level travel demand

Figure 3.10(a) shows the TNC passenger pick-ups for each TAZs on a typical Monday at 9 AM. The rest of the dataset is similar to 3.10(a), except for the time (24 hours in a seven-day week) and events (pick-ups or drop-offs). The hourly TNC pick-ups and drop-offs are scaled to obtain the all intra-SF vehicular travel demand for each hour. The all vehicular travel demand include trips made by not only the TNCs, but also taxis, private cars and public transit vehicles [213]. Table 3.3 shows the estimated ratios between the numbers of TNC and of all vehicular trips by time periods and supervisorial districts (see Figure 3.10(b)) according to the data provided by SFCTA [213]. Contrary to the general traffic behaviour, the proportions of TNC trips are lower during peak hours, which may suggest that commuter trips during peak hours are still primarily taken by private cars. Also, the TNC trip ratios are higher in downtown, indicating better coverage of TNC services in these areas. TNC pick-ups and drop-offs as shown in Figure 3.10(a) are scaled by the TNC trip ratios in Table 3.3 to obtain the origin and destination counts by the hour and TAZ for all vehicular traffic. For example, on average there are 11.8 TNC pick-ups in TAZ 621 in downtown SF on a typical Monday at 9 AM (circled in Figure 3.10(a)). Since it belongs to supervisorial district 6 (circled in Figure 3.10(b)), the ratio between TNC traffic and all vehicular traffic is estimated to be 32% at that hour. So the number of all vehicular trips leaving TAZ 621 on Monday at 9 AM is calculated as 11.8/0.32 = 36.9.

The next step is to connect the origin zones to destination zones and to produce a list of zonal origin-destination (OD) pairs. Like many other travel surveys, the TNC dataset only tells the total number of trips starting (ending) in each zone, but not the corresponding destination (origin) information. In other words, it is known from the TNC dataset that how many people depart from a particular TAZ zone in each hour, but their destinations are not known. To build the complete trip-level model inputs with origins and destinations, a random sampling-based OD matching scheme is adopted. After scaling up the TNC dataset to obtain the total vehicular counts, there are now the "hourly departure vehicle counts"

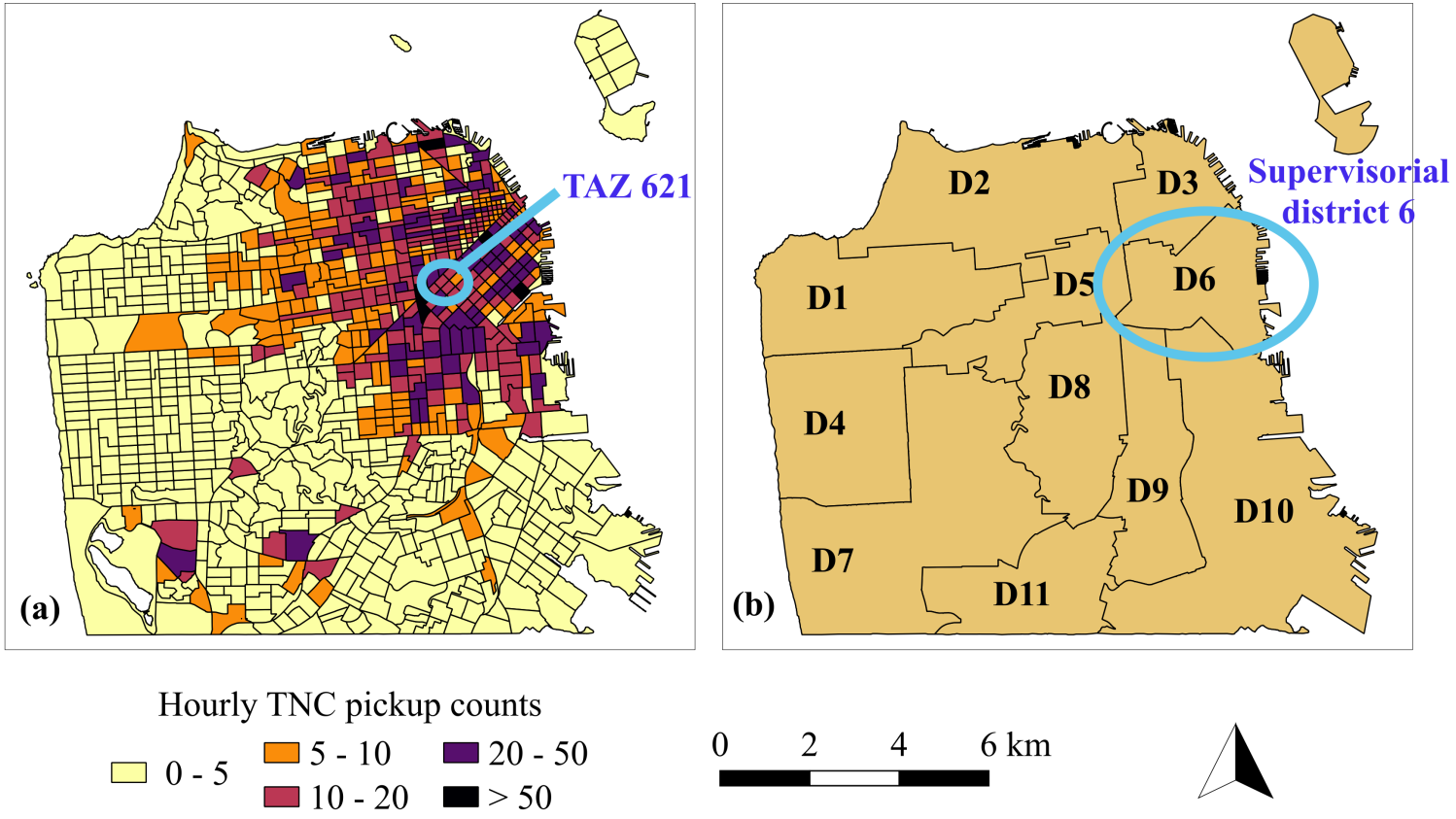


Fig. 3.10 (a) TNC pick-ups on Monday at 9 AM. (b) SF supervisorial districts (reproduced from City and County of San Francisco [50]).

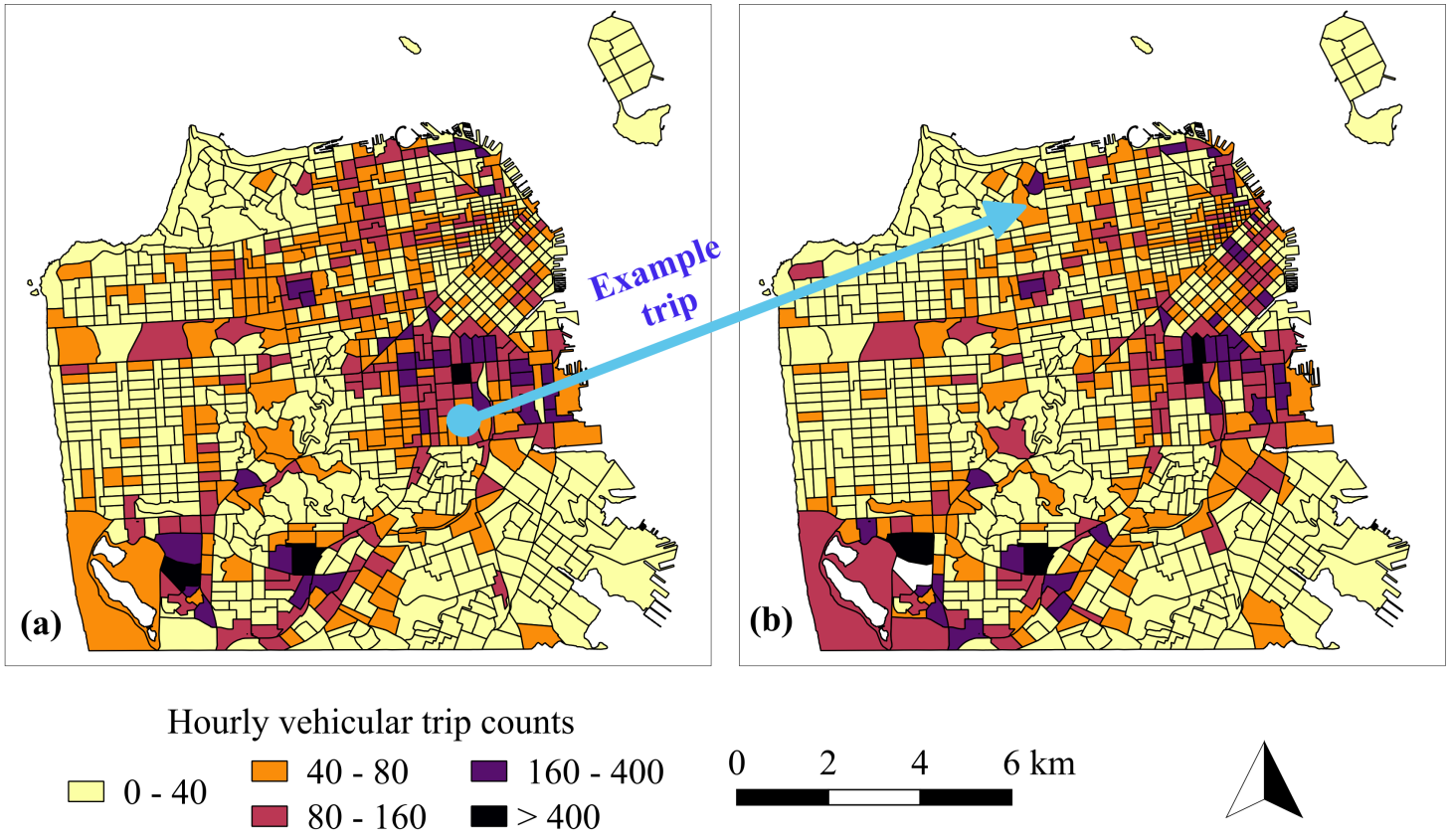


Fig. 3.11 Generating zone-to-zone travel demand through sampling for a typical Monday at 9 AM. (a) Sampling an origin TAZ based on probability defined by zonal-level origin counts; (b) Sampling a destination TAZ based on probability defined by zonal-level destination counts.

Table 3.3 Percentages of TNC trips in all vehicle trips by time periods and supervisorial districts (reproduced from SFCTA [213])

Supervisorial district ID	AM	PM	Off-peak
D1	8%	7%	10%
D2	20%	17%	22%
D3	19%	20%	25%
D4	4%	3%	5%
D5	14%	13%	19%
D6	25%	26%	32%
D7	5%	4%	5%
D8	10%	8%	12%
D9	10%	9%	12%
D10	7%	7%	9%
D11	3%	2%	4%

and "hourly arrival vehicle counts" of each TAZ (these two values are not identical but very close). Such information is used twice. Firstly, the SF hourly travel demand is calculated by taking the sum of the average of the "hourly departure/arrival vehicle counts" of each TAZ. Secondly, the "hourly departure/arrival vehicle counts" of each TAZ are normalised to obtain an estimation of the probability that a trip starts/ends in each of the 981 TAZs. After these preparation, a list of TAZ pairs is generated. The former/latter element in each generated TAZ pair is sampled from the list of 981 TAZs, with the probability of being selected equalling to the normalised "hourly departure/arrival vehicle counts". Short trips, defined as the TAZ pairs with the distances between the starting and ending zonal centroids shorter than 2.5 km (about 30 minutes walking), are excluded as they are considered walkable or bikeable [196]. In the end, zonal-level trip ODs are produced for each hour during the study period of a typical week (Figure 3.11). The total number of trips in each hour equals to the SF hourly travel demand [213].

In the last step, nodal-level OD pairs are generated by selecting a random node within the starting or ending TAZ for each zonal-level OD pair. This leads to a list of node-to-node travel demand (Figure 3.12).

3.4.3 Intercity travel demand

SF is connected to nearby cities through four “gates”: (1) Golden Gate Bridge (to the North); (2) Bay Bridge (to cities on the East); (3) California State Route 1 (SR 1, to the South) and (4) US 101 south and I-280 South (to the San Francisco International Airport, SFO (Figure

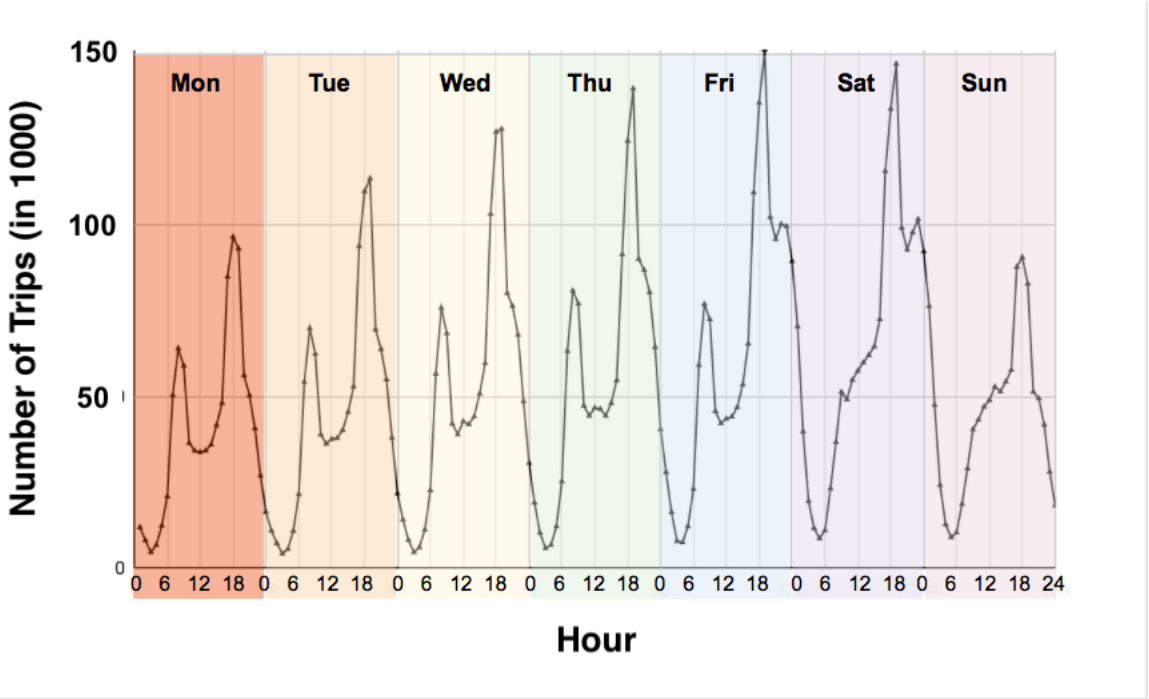


Fig. 3.12 Generated node-to-node travel demand by hour and day of week. The gradual increasing trend of the travel demand from Monday to Friday and the lack of the morning peaks during weekends are the results according to the information in the TNC data.

3.13). So the network can be cut off from the rest of the Bay Area, as has been done in the work of Sall et al. [202]. As the intra-SF traffic was collected during mid- November to mid-December, intercity traffic through the above four "gates" is also obtained for the same period (Appendix 1). It is further assumed that vehicles enter/leave SF via these four "gates" are through-traffic, i.e., they do not start or stop within the city. This assumption can be relaxed in the future by incorporating regional travel surveys (e.g., the California Household Travel Survey, 2010-12) to decide the origins and destinations inside SF for intercity trips.

Table 3.4 shows the daily intercity traffic volumes between the four "gates" in the form of an OD matrix. The numbers in brackets in the last row and column of Table 3.4 are the approximate daily entrances/exits at each of the four "gates" published on their websites. For the Golden Gate Bridge and the Bay Bridge, the daily traffic volumes are not provided by direction, so it is assumed that the traffic is equally split between the entering and exiting directions at these two locations. For the SR 1 and SFO, daily traffic volumes for each direction are known. However, as the differences in traffic volumes between the two directions are close, a rounded average number is taken as the total daily entrances as well as exits for each of them. A matching process similar to the one described in Section 3.4.2 is utilised to obtain a list of intercity OD pairs, based on the observed total entrances and

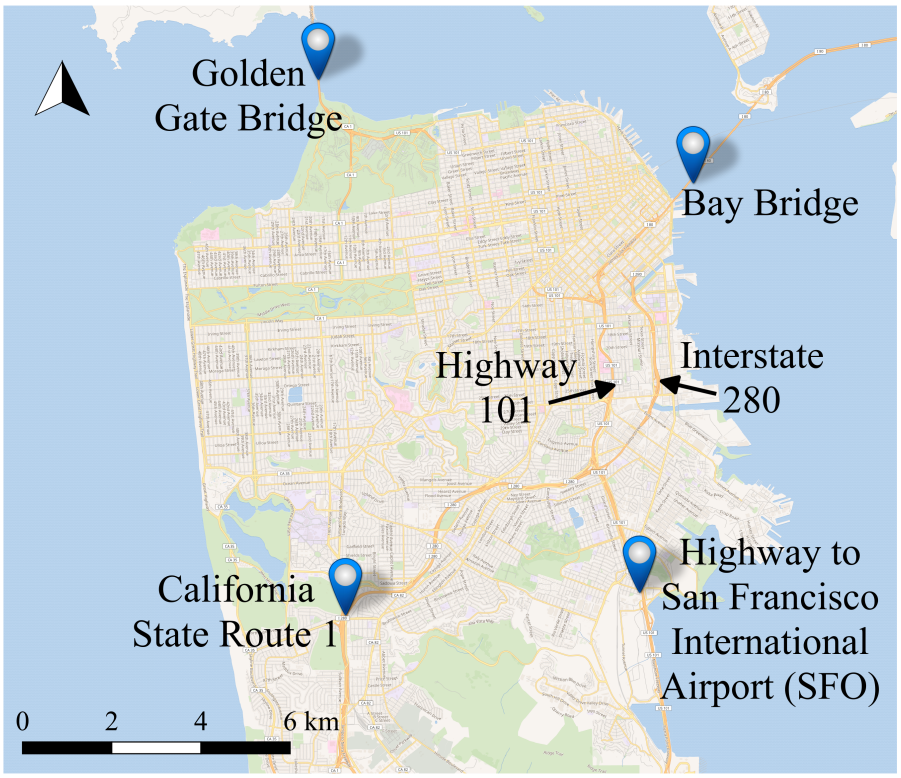


Fig. 3.13 Main entrances and exits to SF.

exits (numbers in the brackets in the last row and column in Table 3.4). The probability of being selected as an entrance or an exit point is equal to the normalised observed daily traffic volume. There are in total 405,000 intercity trips, distributed between the four gates as shown in Table 3.4.

3.5 Model

The intra-SF and inter-city travel demand is loaded/assigned onto the SF road network graph to calculate the traffic distributions during the modelling phase. Specifically, a trip is created for each OD pair and a traveller/vehicle is assumed that traverses the network through a series of graph edges that form the optimum route between the origin and destination nodes. Temporally, the traffic simulation progresses by one-hour time steps. Inside each time step, trips are dispatched in 20 sub-steps and link-level travel times are updated after each batch. The framework for traffic simulation is shown in Figure 3.14.

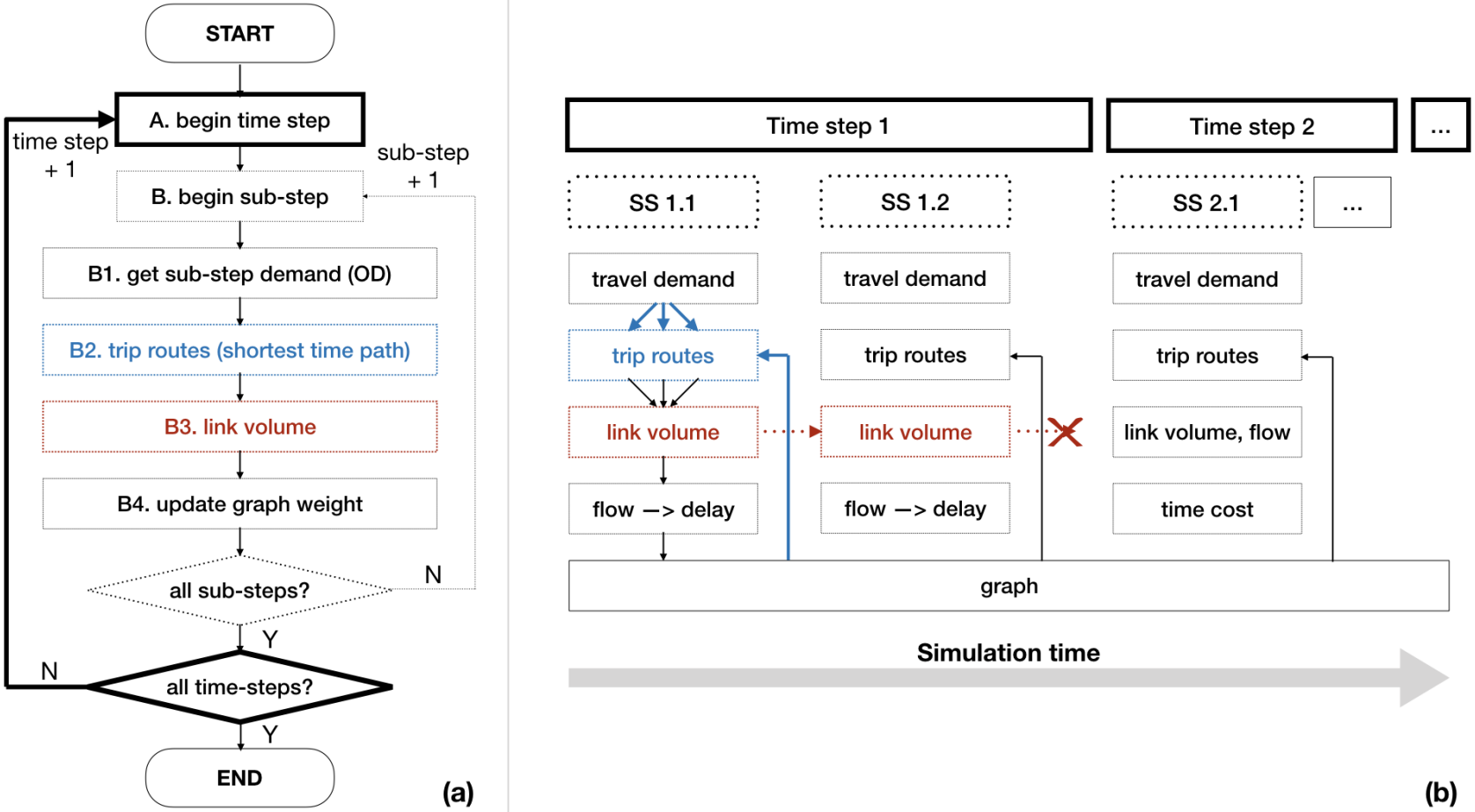


Fig. 3.14 Framework for the traffic simulation. (a) Flow chart; (b) Procedures on the time axis. Meanings of graph elements (1) SS: sub-step. (2) Solid arrow: passing data from the upstream box to the downstream box. (3) Dashed arrow: information (i.e., link-level traffic volume) cumulatively added and passed to the downstream box. (4) Cross: termination of information accumulation.

Table 3.4 Daily intercity travel demand in SF through four "gates"

IN \OUT	GGB	BB	SR 1	SFO	Total IN, model (data)
GGB [1]	-	35,038	19,656	13,895	68,589 (55,000)
BB [2]	35,466	-	64,187	44,703	144,356 (180,000)
SR [3]	19,950	63,433	-	24,974	108,357 (100,000)
SFO [4]	13,926	44,885	24,846	-	83,657 (70,000)
Total Out, model (data)	69,342 (55,000)	143,356 (180,000)	108,689 (100,000)	83,572 (70,000)	404,956 (405,000)

Note: [1] GGB: Golden Gate Bridge. [2] BB: Bay Bridge. [3] SR 1: California State Route 1. SFO: San Francisco International Airport.

3.5.1 The outer loop A: temporal evolution by time step

The outer loop of the traffic simulation controls the progress of time. The body of the outer loop specifies how traffic is simulated within each one-hour time slice that will be explained in detail in the next subsection. For each day simulated, the outer loop is executed for 24 times, starting from 3 AM in the morning (when the traffic is the lightest) till 3 AM in the next day, so the traffic simulation in the first time step of each day can be assumed to run with all link-level travel times equal to their free-flow travel times. In this study, traffic is simulated for the seven typical days in a week.

Essentially, a STA procedure is carried out in each time step. The basic assumption associated with this STA is that all journeys will be finished by the end of the time step. The time step can theoretically be any length, but one-hour is a good choice for the study area as in reality most of the journeys, even during peak hours, can be completed within one hour in SF. As shown in Figure 3.14, traffic conditions from a previous time step will influence that of the next time step through the share of the network graph and edge weights. Specifically, the route choice or traffic assignment of the first sub-step in each time step is directly influenced by the traffic assignment of the previous time step. For example, at the first sub-step of the whole simulation (the first sub-step of the 3 AM time step), it is assumed that all roads are nearly empty and vehicles can move at the free flow speed. But for trips starting at 4 AM and later hours, the travellers will notice that there are already some traffic on the streets. Due to this existing traffic, they can no longer travel at the free flow speed.

In fact, some roads will be more congested than the others, which prompts the travellers to select another route to minimise the journey time.

3.5.2 The inner loop B: traffic assignment in each hour

The inner loop of the traffic simulation implements a routing algorithm to find the route/path for each of the vehicular trip that is scheduled to travel in the specific hour (set by the outer loop). A route or path is a sequence of graph edges that connect the origin node to the destination node. It is assumed in this chapter that a traveller will only seek the fastest route when travelling, ignoring other factors such as route distance, toll cost or habitual route. This is equivalent to (1) setting the time to traverse each road link as the edge weight in the road network graph; (2) finding and assigning the trip to the path that has the smallest total weight (shortest travel time); (3) updating the link-level travel time based on the simulated traffic flow.

Sub-steps are adopted for traffic assignment within each hourly time step, so that the hundreds of thousands of trips belonging to each time step can be assigned in parallel incrementally [164]. Graph weights are updated at the end of each sub-step after each parallel assignment. This has the benefit of stabilising the traffic distribution, avoiding to assign all trips in the hourly time step to a few concentrated path. While it is also feasible to set the number of sub-steps equal to the number of trips (updating the graph after the assignment of each trip), this will make the parallel computing difficult. The sub-step is not a time related concept. Vehicles/trips that are being assigned in any sub-step will persist till the end of the time step, thus link volume will cumulate across the sub-steps till the completion of one time step. The incremental assignment process of each sub-step will be introduced according to the sequence shown in Figure 3.14.

B1. Sub-step demand

The hourly travel demand, expressed in terms of OD pairs, is divided randomly into 20 equal groups, corresponding to the 20 sub-steps used in the simulation. For example, there are around 87,000 trips that are being taken on a typical Monday at 8 AM (63,000 intra-SF and 24,000 intercity). So 5% of the total hourly demand, e.g., 4,350 trips, are going to be assigned (onto the fastest) routes in each sub-step. The random division will bring random outcomes, whose variability will be investigated in the next chapter. Miller [154] uses a similar incremental assignment procedure, but only divided an one-hour time step into 4 increments. The aim of sub-step is to set a fixed frequency for updating the link-level travel time, while allowing all trips within a sub-step to assigned in parallel. When the sub-step

number is too small (e.g., only one sub-step per time step), all trips will tend to choose the same path as the travel time is outdated. However, when the sub-step number is too large, the parallel trip assignment cannot be exploited efficiently. The number of 20 sub-steps per time step is determined through some trial experiments, it was found that 20 sub-steps is sufficient to stabilize the traffic assignment outcomes at acceptable computational cost.

B2. Trip routes

After the sub-step demand has been determined, a priority-queue based Dijkstra's shortest path algorithm library is utilised to find the fastest route for each OD pair (a trip) [132]. Dijkstra's algorithm is an efficient way to find the shortest path between two nodes (vertices), or to produce the shortest path tree from a "source" node to all other nodes in a weighted graph [67]. In the simulation, the Dijkstra's algorithm is applied to the road network graph weighted by link-level travel time and finds the weighted shortest path (i.e., fastest route to travel) from the trip's origin node to its destination node. Dijkstra's algorithm can be implemented with a min-priority queue data structure. The priority queue stores all unvisited graph vertices (whose shortest distances to the origin vertex have not been decided), prioritised according to their current (not necessarily the shortest) distances to the origin vertex. This implementation allows the Dijkstra's algorithm to take advantage of the fast operations of the priority queue, including inserting a node, extracting the node with the lowest priority and decreasing the priority of a node [157].

The Dijkstra's shortest path algorithm is executed for each OD pair in the traffic simulation at every time-step. This quickly becomes the computational bottleneck due to the large number of trips and the large graph size in a city-scale simulation. However, as the graph weights do not change within a sub-step of the incremental assignment process (Figure 3.14), the choice of route for one trip is not affected by the route assignment of the other trips in the same sub-step. This is the reason that the Dijkstra's algorithm can be executed in parallel for trips in the same sub-step.

In this study, the SF model is deployed on a single-node of the Cambridge HPC cluster CSD3 (https://www.hpc.cam.ac.uk), utilising 32 cores (parallel processes). The hourly number of trips to be computed varies from 4,111 (Tuesday at 3 AM) to 151,170 (Friday at 7 PM). The number of trips in a sub-step (5% of hourly total trips) is between 200 to 7,500. This adds up to 9 million trips for the simulation of a typical week. It took around 40 minutes to simulate the traffic of a typical week in the SF road network with 9643 vertices and 26893 edges.

In this chapter, it is assumed that all travellers are fully aware of the travel time of each road link at the time of departure. In reality, this reflects a situation where the traffic condition

of each road is continuously monitored by traffic sensors or crowd sourcing data from e.g., GPS devices or cell phone signals. This assumption will be investigated in the next chapter by downgrading the sensor coverage and revealing a mixture of real-time and delayed traffic condition information.

B3. Link volume and equivalent hourly flow rate

Link volume is the number of vehicles that have traversed a link in each time step. The link volume can be considered as an automatic traffic counter that keeps track of how many vehicles have passed a link in each hour. In the sub-stepping process described in this chapter, the link volume always initialise with zero at the first sub-step of each time step, assuming all travellers from the previous time step have finished their journeys. Starting from the second sub-step, the volume counter will report the cumulative numbers of vehicles that have been assigned to each link.

The link volume is essential to lead to an estimation of the hourly flow rate. Flow rate is defined as the number of vehicles that pass a link in a unit time, which is assumed as one hour in this study. As there is no time associated with sub-steps, it is impossible to get the flow rate according to the standard Equation 3.2a. Instead, the cumulative sub-step volume will be scaled according to the OD pairs that have been assigned out of all the OD pairs in an hourly time step, as shown in Equation 3.2b.

$$flow\ rate = \frac{volume}{time\ duration} \tag{3.2a}$$

$$flow\ rate = \frac{volume}{demand\ assigned} \times hourly\ demand \tag{3.2b}$$

B4. Link-level travel time update

Link-level travel time is set according to the well-known Bureau of Public Roads [34] volume-delay curves. It has the following form [52]:

$$t_t = t_f \times (1 + \alpha(\frac{q}{capacity})^{\beta}) \times f_p \tag{3.3}$$

where t_t is the link-level travel time with traffic on the link; t_f is the free flow travel time of the link; $capacity$ refers to the link capacity and as t_f are model inputs determined in Section 3.4. α, β are calibration parameters set to 0.6 and 4 and f_p is the city-specific correction factor (1.3 for SF based on [52]). q is the hourly flow rate and is initialised as 0 in the first sub-step of each time step, or updated as the sum of vehicles assigned to a particular link in all previous sub-steps. It is acknowledged that the values of the calibration parameters may

have big impacts on the final outcomes. For example, a larger β means that drivers are more sensetive to congestions (time penalty due to high traffic volume is higher), which would encourage people to take longer but less congested routes. The adopted values are those that are the most commonly used in the literature.

Unlike the link-level volume that is reset to 0 at the beginning of each time step, link-level travel time at the beginning of each time step is set according to the simulated traffic from the previous hour to ensure some continuity of the congestions. Then after each sub-step, a new link-level travel time is calculated according to the newly updated the flow rate. This updated link-level travel time will be fed to the next sub-step to continue the process of the traffic simulation.

3.6 Results

A direct result from the traffic simulation is the traffic volume on every road link in each hourly time step. Figure 3.15 shows such results on a typical Friday at 6 AM (off-peak) and at 6 PM (evening peak hour). At 6 AM, the traffic is light on most roads in the city except for the highways. While at 6 PM, significantly more traffic is seen on the highways as well as roads in the inner part of the city. The simulation results indicate that during the evening peak hours the traffic condition on Highway 101 is worse than on Interstate 280 (I280, route number shown on Figure 3.13), which is in agreement with the local knowledge. However, downtown SF (in the northeast) is not particularly congested. This is probably because of the dense placements of traffic lights in the city centre, making it more delayed and thus less attractive to go through these streets in the journey.

Figure 3.16 shows the volume-to-capacity ratio of each road link on a typical Friday at 6 AM and at 6 PM. It is obtained by dividing the traffic volume data in Figure 3.15 by the capacity of each road as calculated Section 3.4. At 6 AM, traffic is below capacity on almost every road. While at 6 PM, there are 9.2% of the roads have traffic that exceeds capacity. The critical links with high volume-to-capacity ratio are not exactly the same as the links that have more traffic. For example, the Bay Bridge (the long link in the northeast corner) has almost the heaviest traffic flow in SF at Friday 6 PM. But as it also has a high capacity, its volume-to-capacity ratio is only 1.3.

The results are also displayed in terms of relative delay, the ratio between the travel time in traffic and the free flow travel time. Figure 3.17 shows the time delays of three street segments for a five day week. The three streets have different road classes but the patterns of morning and evening peak are clearly seen on the two busy ones. This pattern follows hourly demand in Figure 3.12. The large delays on key routes indicate potential to reduce

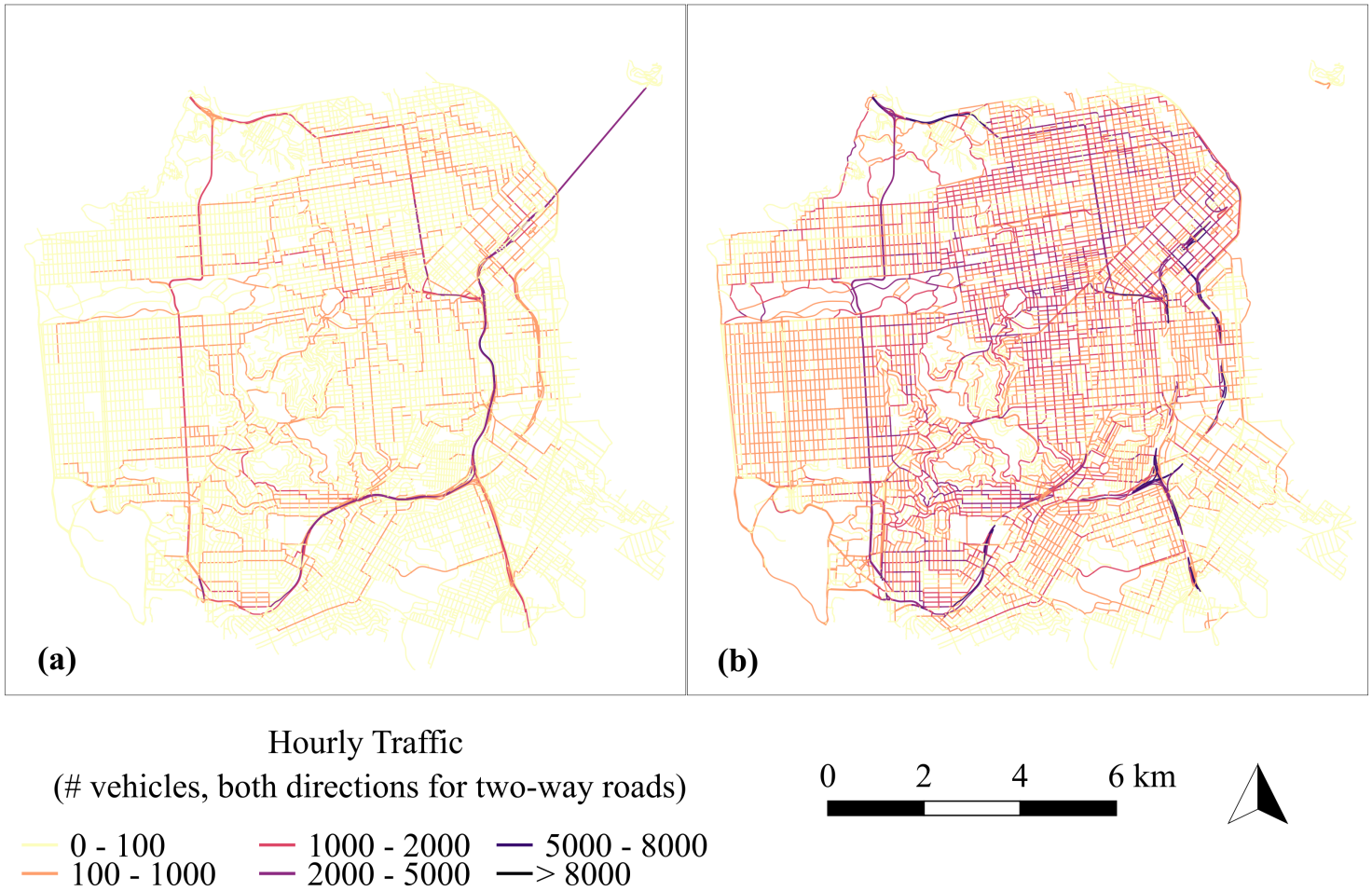


Fig. 3.15 Hourly link-level traffic volume from the mesoscopic traffic simulation. (a) Friday at 6 AM. (b) Friday at 6 PM.

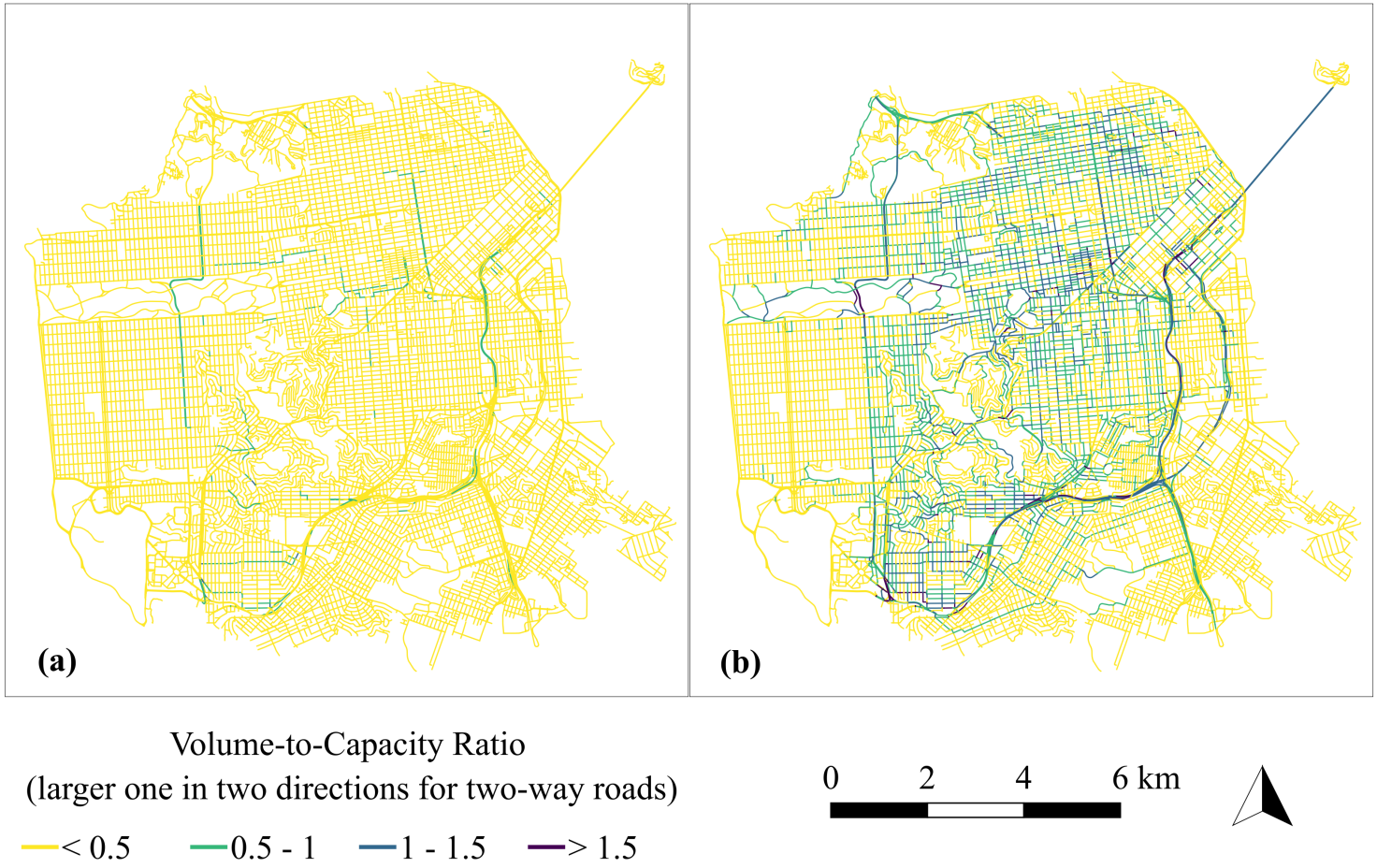


Fig. 3.16 Link-level volume-to-capacity ratio. (a) Friday at 6 AM. (b) Friday at 6 PM.

congestions during peak hours and improve the traffic network efficiency with, e.g., flexible working hours.

As no systematic real traffic volume observations have been found for the city of SF, the simulation results are instead compared with the outcomes from another existing model as a preliminary validation step. Figures 3.18 compares the results from the mesoscopic simulation in this study with results from the Bay Area MTC Travel Model One outputs [161]. Specifically, the "2015_06_022" scenario in the MTC model result repository is chosen for comparison as its forecast year of 2015 is the closest to the year of 2016, when the travel demand data used in this study were collected. The MTC model utilises a simplified road network for the whole Bay Area but predicts the traffic distributions of different types of vehicles (passenger cars, commercial vehicles of various size, etc.). Temporally, it provides results for five time periods on a typical weekday (early morning, AM peak, midday, PM peak and evening). As weekends are not covered by the MTC model, it is decided that the total weekday traffic from the two models should be compared. Figure 3.18(a) shows the total weekday traffic (Monday to Friday, 24 hourly time steps per day) from the mesoscopic simulation in this study. Figure 3.18(b) is the counterpart from the MTC's model, cropped to the study area.

An initial visual inspection of Figure 3.18 suggests that similar total weekday traffic is predicted by these two simulation models. However, it is difficult to conduct more quantitative comparisons of the results from Figure 3.18. To begin with, there is no common field in these two networks to identify corresponding road links, such as street names or the road ID. Besides, the MTC network is more sparse geographically and has more simplified road geometries than the OSM data, so the same roads do not align spatially in these two maps.

Due to these practical difficulties in aligning the results from these two models, an alternative approach is taken to facilitate the comparison instead. First, the road links in each model are reordered, ranked by the total weekday traffic in descending order. Here, the "total weekday traffic" is used as a measure/proxy of link criticality. The more traffic is on a link, the higher its criticality and the higher its rank among all links are. Based on this criticality rank, a "mileage" number is calculated for each link by taking the cumulative sum of link length (total length for roads ranked before itself).

Road links ranked higher receive smaller mileage numbers. Figure 3.19 shows the scatter plot of the link-level weekday traffic against the link mileage. It can be seen that (1) the total length of the MTC network (1,328 km) is about 40% of the OSM network (3,142 km), which is expected as the MTC network contains only important roads and leaves out many residential streets, while the OSM network includes all drivable roads; (2) the data series associated with the mesoscopic model results in this study (labelled as "SF mesoscopic" in

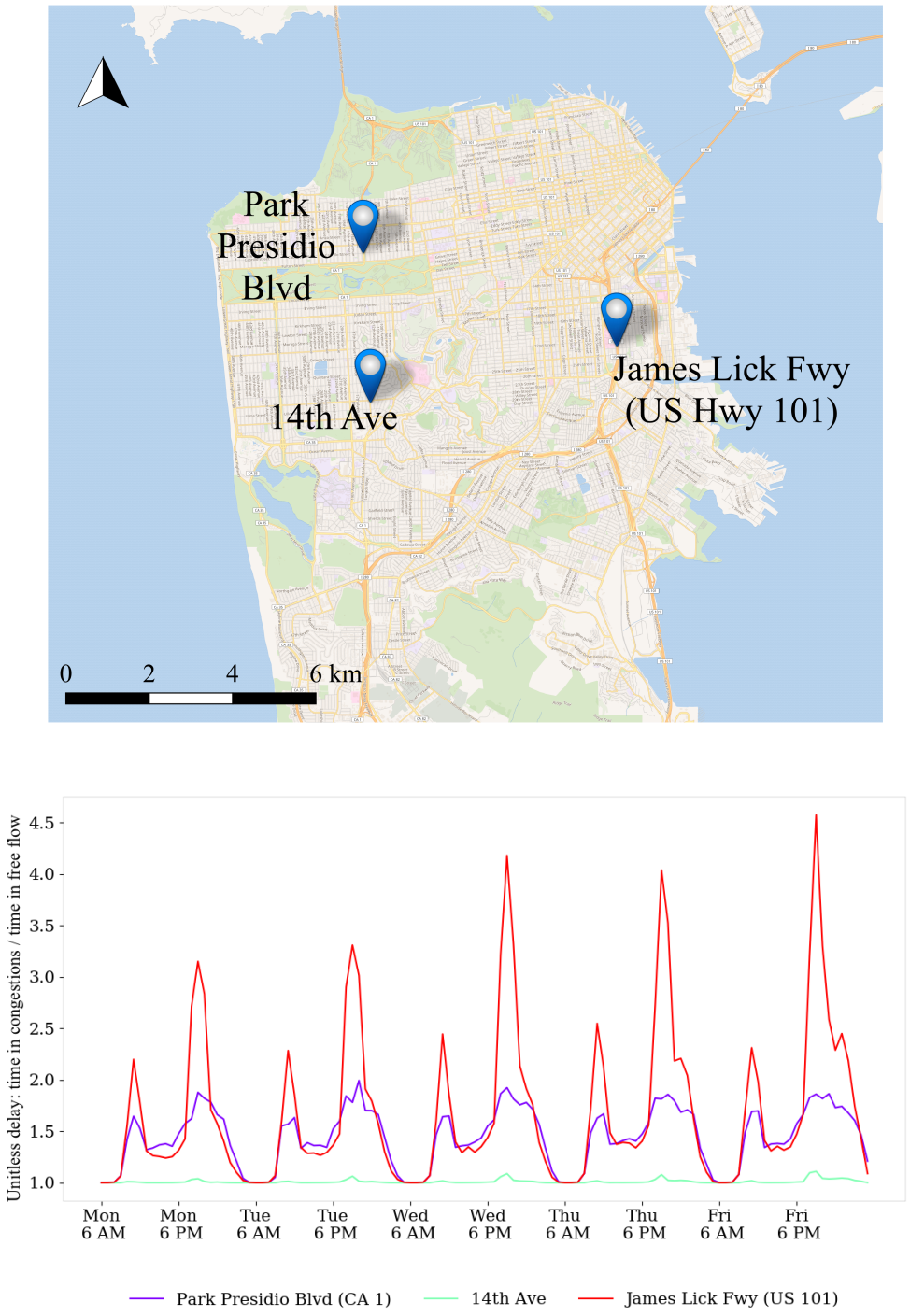


Fig. 3.17 Time delay for three roads in each hour of a week.

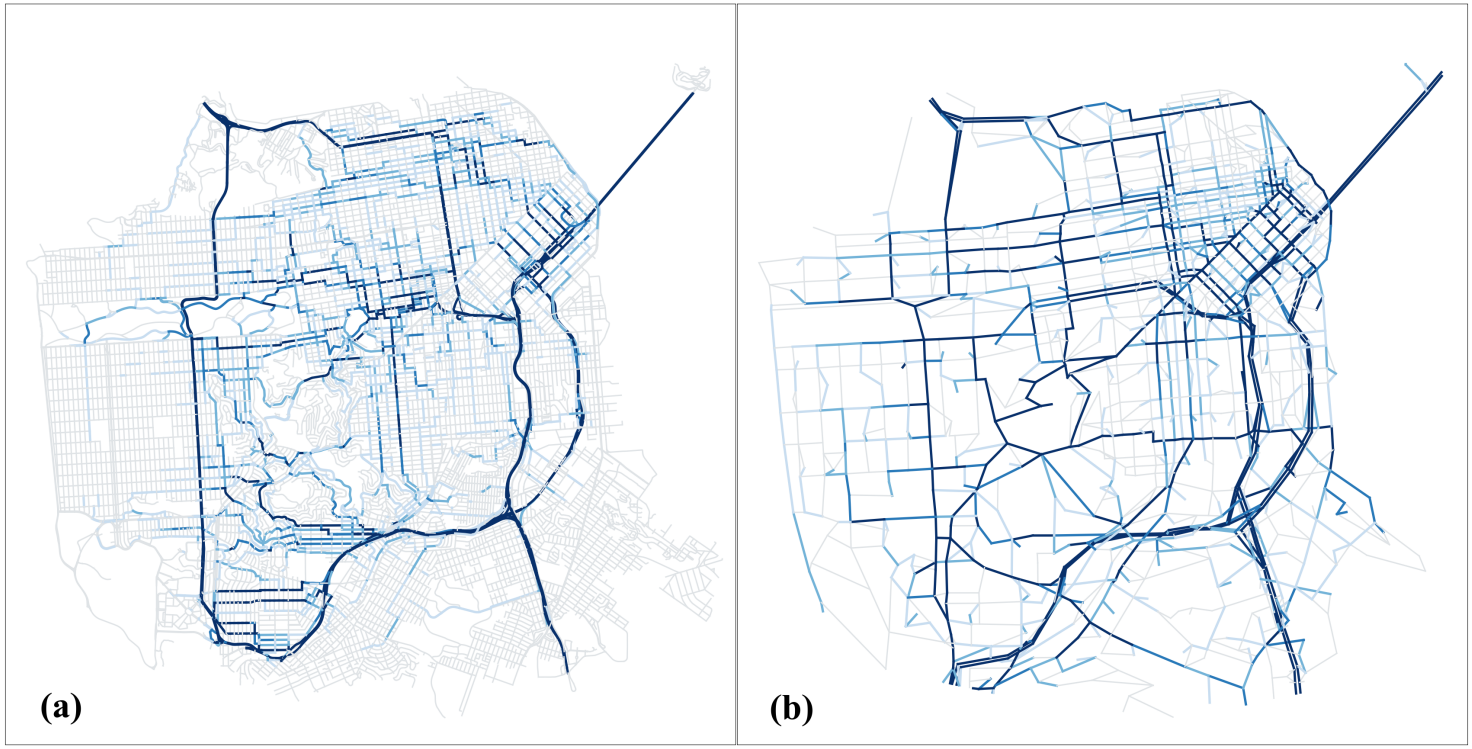


Fig. 3.18 Total traffic volume in SF from Monday to Friday from two simulations. (a) Simulation results from this chapter. (b) MTC Travel Model One (reproduced based on data from [161]).

Figure 3.19) is higher when the road mileage is below 100 km, indicating that the mesoscopic traffic model this study predicts more traffic than the MTC model for the top 100 km of roads ranked in terms of traffic volumes (the heavier traffic regime of the network); (3) the two data series overlap after 100 km of mileage, indicating that similar amounts of traffic are predicted by the two models for road links with mileages between 100-1000 km.

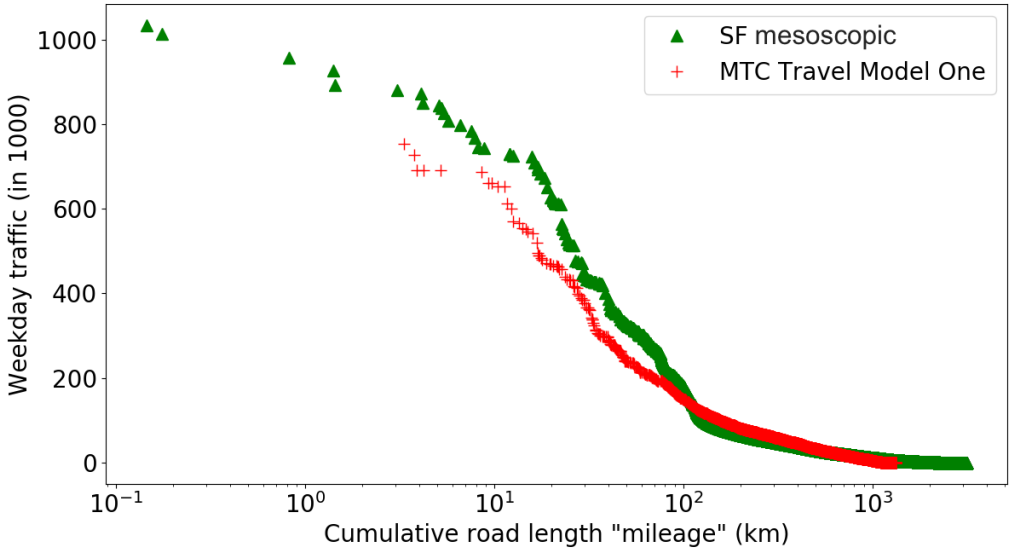


Fig. 3.19 Weekday link-level traffic against cumulative link length (“mileage”).

A similar comparison can also be generated for the link-level travel time in congested situations. Figure 3.20 shows the plot of link-level travel time against the cumulative mileage. Specifically, for the MTC data, the link traversal time during the evening peak period is used. For the mesoscopic traffic simulation developed in this chapter, the link traversal time at 6 PM is plotted for each weekday. The comparison between these two models shows similar trends as in Figure 3.19. When the mileage is small (below 30 km), the link traversal time results from this study are higher than the MTC results, indicating again that the mesoscopic traffic model predicts heavier traffic in this regime. When the mileage is higher than 30 km, the differences between the two model results are less obvious.

3.7 Summary

In this chapter, the development of a mesoscopic traffic simulation model is presented. The mesoscopic model combines the macroscopic relationships between link-level traffic parameters (the BPR volume-delay relationship) with finely disaggregated trip-based travel

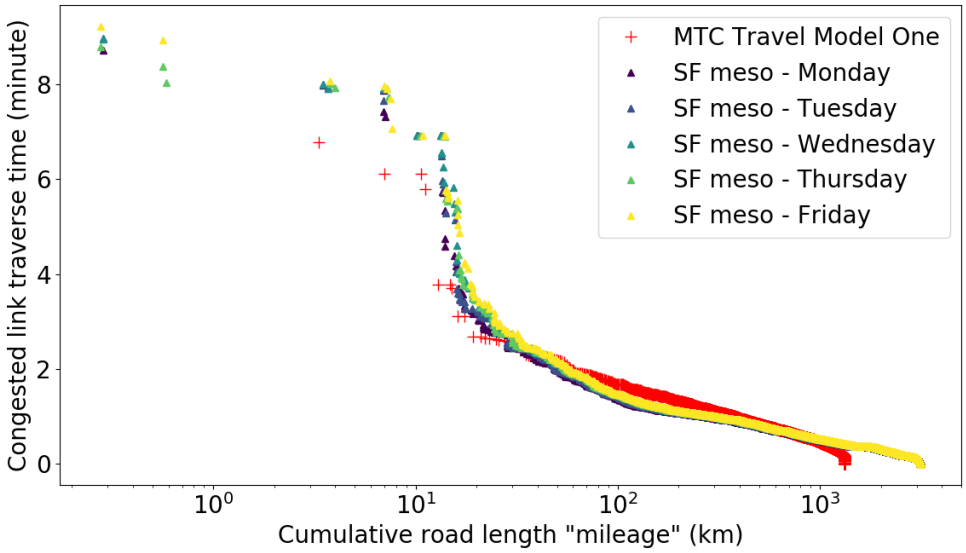


Fig. 3.20 Link-level travel time in congested situations against cumulative link length ("mileage").

demands. The road network graph used for the simulation is obtained and processed from the OSM, with some corrections such as removing redundant nodes and calculating the gradient adjusted link capacity. The travel demand data consist of two parts: the first part is the intra-SF trips scaled up from the TNC data by the SFCTA; The second part is the inter-city travel demand collected at four "gates" at the border of SF.

The traffic simulation model is essentially STA with hour-long time slices. To stabilise the traffic distributions, 20 sub-steps are adopted where the hourly travel demand is assigned incrementally. Within each sub-step, the fastest routes are calculated in parallel using the Dijkstra's shortest path algorithm. At the end of each sub-step, the new link-level travel times are calculated based on the cumulative link-level volume, which will be used to guide the route choices of the next time-step. With these implementations, traffic simulations on the SF road network (about 10,000 nodes and 25,000 links) for a seven day week can be calculated in less than one hour.

The results from the traffic simulation include, e.g., hourly traffic volume distributions, volume-to-capacity ratios and link-level travel times. These results indicate that the mesoscopic model can capture the spatially and temporally varying dynamics of the city-scale traffic for each hour and each road link. To validate the simulation, weekday traffic volume and link traversal time results are compared with those obtained from an official model. The comparison suggests that the mesoscopic traffic model is able to generate similar traffic

distribution patterns as the official model. However, in future studies, it is still desirable to validate the simulation outputs against real observations.

Chapter 4

Value of live information on traffic system efficiency

4.1 Overview

In the traffic simulation model described in Chapter 3, it is assumed that all travellers have access to the up-to-date traffic condition information when making a journey route choice. This assumption may not be far from the reality in San Francisco (SF), where the smart phone coverage and ownership rates are high. In fact, it is among the first few cities that Google Maps rolled out its live traffic information service in 2007 (Figure 4.1). However, it is still unclear how this assumption affects model results. In this chapter, more simulation experiments will be conducted aiming to quantify the effects of relaxing the assumption of perfect information on the traffic simulation results.

Specifically, the source of real-time information in this chapter is assumed to come from mobile Global Positioning System (GPS) devices, as adopted for traffic condition estimation in most mobile navigation applications. Certain percentages of the vehicles in the traffic simulation are assumed to have GPS-enabled mobile devices and can collect and report their travel times of each road link (they are called probe vehicles). As each individual driver may experience slightly different travel time on the same link, due to events such as traffic signals or timid or aggressive driving behaviours, a stochastic variation term is added to the data reported by each probe vehicle. Note that as the GPS speed sensor itself is usually accurate, the difference in individual link traversal time is termed "variability" rather than "error" throughout the chapter. The probe ratio and probe data variability together are adopted to emulate the live traffic data collection process in the real life. Simulations are run for multiple times under the same set-up (with a fixed level of probe ratio and data variability), so as

to test the influence of the random/stochastic processes involved on the traffic simulation outcomes. In Chapter 3, each driver is assumed to have access to the perfect traffic condition information, i.e., the probe ratio is assumed to be 100% with no variability. Such assumptions in Chapter 3 are relaxed in this chapter by introducing a larger range of probe ratios and probe data variabilities.

Randomness and stochasticity are two similar concepts. Both lead to variabilities in the outcomes that can be quantified through repeated experiments. In this thesis, a distinction is made between these two concepts as the following: a random process does not involve an underlying distribution and the results of a random process will not show any systematic pattern. An example of such a process is the random split of the total travel demand into sub-steps. While stochasticity is due to the use of variables that follow certain distributions. Repeated experiments of a stochastic process may lead to systematic patterns in the results. An example of the stochastic process would be the use of the *Gamma* distributed probe data variability coefficient in this chapter.

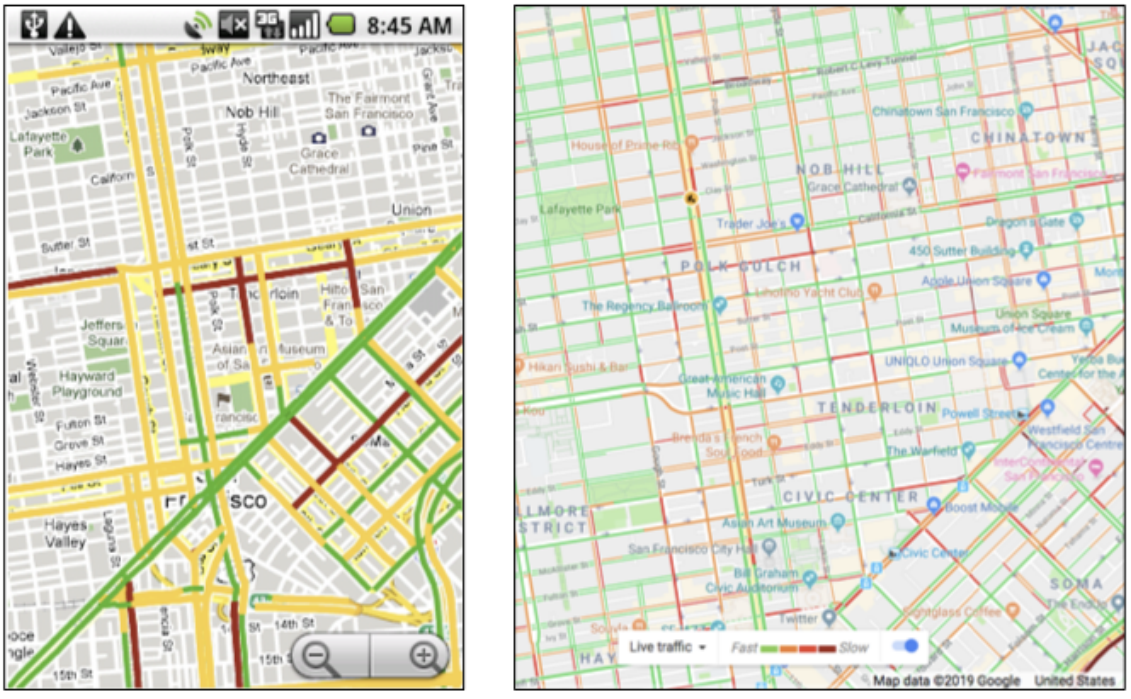


Fig. 4.1 Comparing Google Traffic product in 2009 and 2019. (a) Live traffic coverage is only provided for a few major streets in downtown SF in 2009 [95]; (b) Live traffic coverage is available for all streets in downtown SF in 2019 [96].

4.2 Background

Real-time traffic data are of great benefit and importance to both traffic managers and individual drivers. Traffic managers rely on such information to implement traffic control measures from adaptive signal timing to dynamic congestion charging [234, 258] or monitoring emissions [99]. While for individual drivers, through the Advanced Traveller Information System (ATIS), knowledge of up-to-date traffic condition can help finding alternative routes in congestions, avoiding incidents and reducing errors in travel time estimations [125, 136].

A variety of sensors and techniques has been used for real-time data collection for the above applications. For a long time, the most popular source of traffic information is from the inductive loop detectors buried underneath the pavements, where single-loop detectors can measure flow and lane occupancy while dual-loop detectors ("speed traps") are also able to report speed and vehicle classification in addition [243]. However, loop detectors, as well as other intrusive sensors such as pneumatic road tubes, have major drawbacks due to their limited spatial coverage range, expensive installation and maintenance [250]. On the other hand, non-intrusive technologies provide alternative ways to collect traffic data. In fact, some non-intrusive sensors such as the magnetic, ultrasonic and microwave sensors were adopted before the loop detectors in the 1960s [155]. Other methods belonging to the non-intrusive category such as radio-frequency identification and license plate recognition have also been widely used nowadays for vehicle identification, volume and speed detection in tolled roads, law enforcement and general traffic monitoring. Compared to intrusive sensors, the non-intrusive methods are easier to install and maintain. Yet, they too suffer from disruptions from the environment, high hardware cost and the limited coverage thereafter [250]. In contrast to the fixed intrusive and non-intrusive sensors, the Floating Car Data (FCD) based on probe vehicles are becoming increasingly recognised for their values in traffic sensing. Primarily, it takes advantage of the data collected by existing vehicles in the traffic network (taxis, buses or private vehicles) and no instrumentation is needs to be set up on or along the roadway [63]. The low cost and continuous coverage make the FCD an attractive method for real-time data collection in a modern traffic system.

Before the wide adoption of smart phone devices, FCD mainly come from a small number of dedicated probe vehicles such as taxis or bus fleets equipped with special GPS devices. Lorkowski et al. [143] proposed the system of FCD based on taxi position data and demonstrated its ability in network-level traffic monitoring, automatic congestion detection, dynamic routing and map creation. The system was mainly intended to be used by commercial fleets as it was believed at that time that ordinary drivers would be unwilling to pay for navigation services. Another case of dedicated vehicle for traffic speed monitoring is the Large Scale Floating Car Data (LSFCD) system in Italy, where the entire motorway network

(> 6,000 km) was monitored by 0.6 million vehicles. GPS on board units were installed on these vehicles by car insurers [63].

With the wide spread use of mobile phones, it is now able to source FCD from a much larger population at low costs. This process of gathering, analysing and sharing local knowledge through everyday mobile devices is called participatory sensing. Position and localisation of mobile phones can be obtained through cell-towers, wireless network traces and GPS. Due to the low resolution of cell-tower data and the limited coverage of wireless network traces, GPS enabled mobile phones have become the most sought after method for mining mobility patterns [140]. Feasibility of such systems is confirmed by field experiments, such as the "Mobile Century" project, where real-time speed data were collected by GPS enabled mobile phones carried by more than 100 volunteers driving down a highway in loops, maintaining a fixed 2-5% penetration ratio throughout the test [103]. This further grows into a larger experiment and service called the "Mobile Millennium" from 2008 to 2010, providing a regional-wide traffic monitoring and information application to 2,000 smart phone users [156].

There have also been several theoretical/simulation research studies on mobile phone FCD, mainly investigating their performance on achieving certain tasks, such as speed monitoring. Compared to field experiments, it is easier to explore a wider range of the parameter space in simulations, such as the probe penetration rate, GPS reporting time interval and error rate. In Tao et al. [226], a simulation is conducted on a small network made up of 14 nodes and 22 links, assuming a GPS penetration rate of 10% and location sampling interval of 10 seconds. The emulated GPS location samples are post-processed to provide link-level speed estimations that are dynamically updated every 10 minutes. Another simulation example is from Gayah and Dixit [87], where the feasibility of estimating network-level averaged traffic density is tested based on FCD in a network with 110 intersections. In their simulation, the probe penetration rate varies from 2.5% to 50% and the location sampling interval ranges from 15 to 300 seconds. The network-level averaged traffic density result can be used in conjunction with the Macroscopic Fundamental Diagram (MFD) for a control scheme [256].

With the advent of commercial navigation applications such as Google Maps, Waze, Here, INRIX and TomTom, the scale that public can access real-time traffic speed and incident data becomes unprecedented. According a survey conducted by Pew Research in 2015, nearly one third of Americans rely on smart phones for turn-by-turn navigations while driving [186]. These services not merely have a high penetration rate, but also feature sophisticated algorithms that integrate various location data sources (GPS, Wi-Fi, mobile networks, sensors), filter out abnormalities (e.g., buses stops) and provide accurate and

frequently updated estimations of the current and forecast network status. Commercial navigation deployments together own the largest FCD collecting databases in the world, but at the same time, are prone to privacy and security fallibilities [19, 120].

4.3 Motivation

The primary goal of this chapter is to study the sensitivity of the traffic simulation results with regard to varying levels of traffic information availability and variability. This is to test the assumption in Chapter 3 where every traveller always has access to the most up-to-date traffic information at the time of departure (100% probe penetration rate, no variability in link traversal time measured by different vehicles in the same simulation sub-step). It is important to investigate how this assumption of "perfect information" affects the simulation outcomes. In addition, at each information availability level, a few repeated simulations need to be conducted to understand the range of fluctuations in the simulation results due to the randomness and stochasticity involved.

A practical implication of the study in this chapter is to answer questions such as the minimum probe ratio required for an efficient traffic navigation system. The traffic information collection process in this chapter mimics the services provided by Google Maps and Waze type applications. By enlisting GPS enabled mobile devices, network-wide traffic data can be collected in the most cost effective manner. But in the mean time, the participants are also at the risk of possible privacy issues. The purpose to define a minimum probe ratio is to find a balance point where the system performance can be ensured while minimising the exposure of personal data, for example, by only collecting information from the taxis. The experiments in this chapter are designed accordingly where only a certain percentage of travellers share their location and speed information (probe ratio) while other travellers selfishly utilising such information for route planning without the need to contribute back to the database directly. A *Gamma* distributed stochastic variation term is attached to individual probe measurements to reflect the variabilities in individual behaviours that are not included in the basic traffic simulation framework in Chapter 3. The information gathered by probe vehicles is then aggregated to the link level and revealed to the travellers who depart at the next sub-step.

4.4 Model structure

In this chapter, the structure of the traffic simulation model will be adjusted to reflect the collection and utilisation of real-time traffic speed data by GPS enabled mobile phones as

floating sensor nodes (probes). The basic simulation flow is the same as introduced in Chapter 3, except for including additional steps to process the link-level travel time estimations based on the probe data.

Figure 4.2 illustrates the differences in the simulation process with or without probe vehicle's presence. When a probe vehicle is en-route (e.g., the green coloured GPS enabled vehicle), it records and reports its instantaneous speed at every few seconds. These speed reports are matched to the nearest road links so as to estimate the travel time on the links. Reflected in the simulation model of this chapter, this means that when one or more probe vehicles pass a link, they will be able to report the time to traverse the link. The road network weights, i.e., the link traversal time of each edge, will be updated according to the measured link-level travel time. In addition, based on this reported travel time, albeit with some variations from the true travel time, the hourly flow rate and current link volume can also be estimated by inversely using the Bureau of Public Roads (BPR) relationship (see Equation 3.3 in Section 3.5.2 for an explanation of the BPR volume-delay curve). This knowledge of the inferred volume and flow rate does not need to be utilised immediately, but can be stored in a traffic management system as prior knowledge when no probe vehicle passes next time.

By contrast, when none of the vehicles on the link acts as the probe, there will be no new information as for the current travel time, flow rate or volume on the link. As a result, the travel time has to be calculated based on the best knowledge available, rather than to be measured directly. In this case, the best knowledge available is the knowledge of the congestion level experienced and reported by the most recent probe vehicle, which shall be adjusted to estimate the the current flow rate and link-level travel time. Without any probe, this situation is equivalent to the traffic managers not aware of any increase in traffic and effectively diluting the old link volume by a larger denominator as the simulation sub-step increases, which will result in underestimated flow rate and travel time.

A more systematic depiction of the probing, travel time estimating and updating process integrated in the traffic simulation is shown by the flowchart in Figure 4.3. The steps (e.g., B, C1, C2, ...) are labelled in accordance with the basic framework in Figure 3.14. A description of each step is given as the following.

B. Trip assignment in a sub-step

The box labelled B in Figure 4.3 corresponds to B1 to B3 in Figure 3.14 of the basic framework in Chapter 3. In this step, an optimum path is found for each trip in a sub-step. In the simulation in Chapter 3, it is assumed that all travellers are fully aware of the actual link-level travel time, so the concept of optimum path in Chapter 3 is equivalent to the actual fastest route that leads to the destination from a given origin. While in this chapter, as the

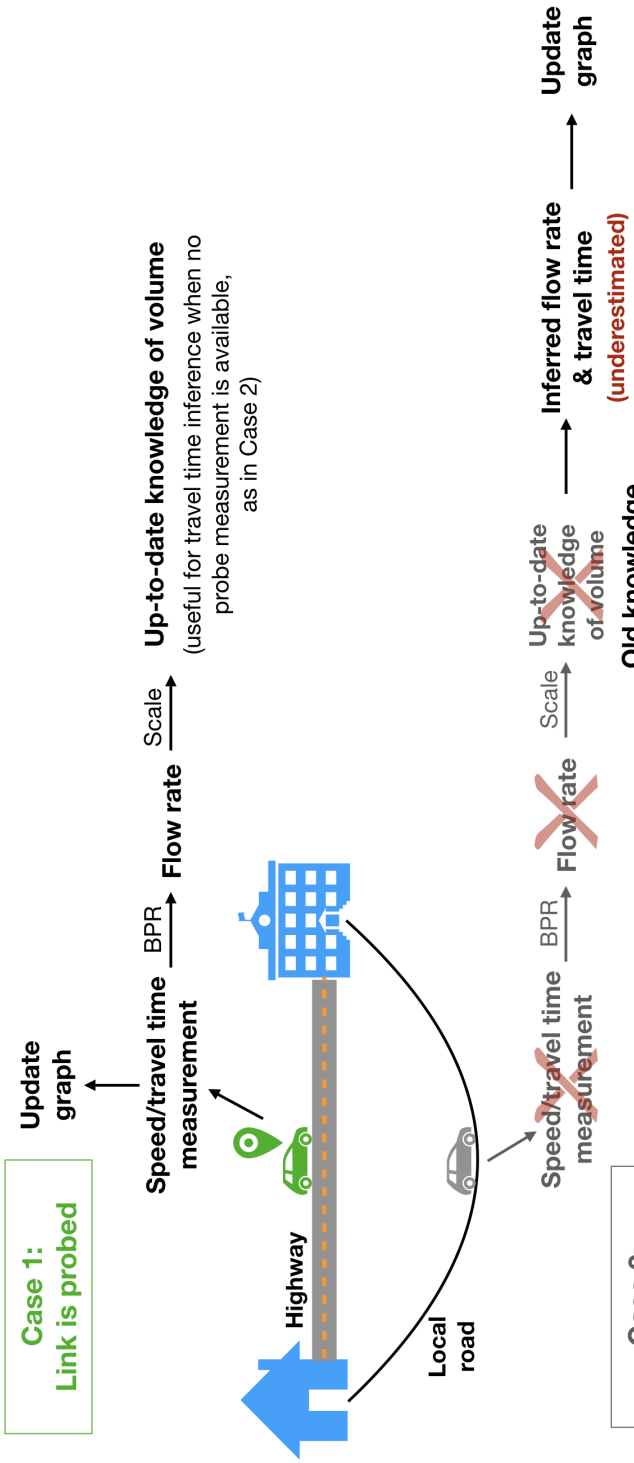


Fig. 4.2 Updating of travel time knowledge based on traffic conditions from probe data.

correct link-level travel time is not always made known, the optimum path becomes the path that is believed to be the fastest route from the origin to the destination at the time of departure. For example, in Figure 4.2, when the highway is congested, the local road may be the actual fastest route to travel from "Home" to "School". While if no real-time traffic condition information is available on either roads, the traveller may still choose to travel on the highway as they think the highway is the faster option. The determination of the "perceived link-level travel time" will be explained in detail in Step C1 (with probe information) and C4 (without probe information).

After assigning the vehicles to their respective optimum paths, the end result of this step is the link-level traffic volume ("true link volume") at this the sub-step. Like in Chapter 3, this true link volume will cumulate across sub-steps within the same time step, reflecting the cumulative traffic counts in each time step. As it is still assumed that all travellers should finish their journeys within a time step (static assignment, whose validity will be shown in the results of this chapter), the cumulative traffic counts will be reset to zero at the end of each time step.

C. Determining which links are probed

In Step B, tens to hundreds to thousands of vehicles are assigned to their optimum paths based on the perceived travel time. Among them, only a portion of vehicles actually act as floating car sensors (probes) and report the live traffic information. This portion, or the probe penetration rate, and the vehicle participation status (probe or non-probe) are specified at the beginning of the simulation. Links that are part of the collection of paths traversed by probe vehicles are "probed", while other links that only carry non-probe vehicles or are not used at all are "not probed". Traffic managers need to update the link-level travel time for each road link at the end of a sub-step. Depending on whether a link is probed or not, two different procedures are taken to calculate the perceived link-level travel time.

C1 & C2. Updating travel time and traffic volume for probed links

For probed links, the probe vehicles report their experienced link-level travel time. This measured travel time can be used directly to update the road network graph for the next sub-step as shown in step C1.

$$t_{t,l} = \frac{1}{N_p} \sum_{p \; in \; N_p} t_{t,l,p} \tag{4.1}$$

$$t_{t,l,p} \sim Distribution(\mu_{t,l}, \sigma^2_{t,l}) \tag{4.2}$$

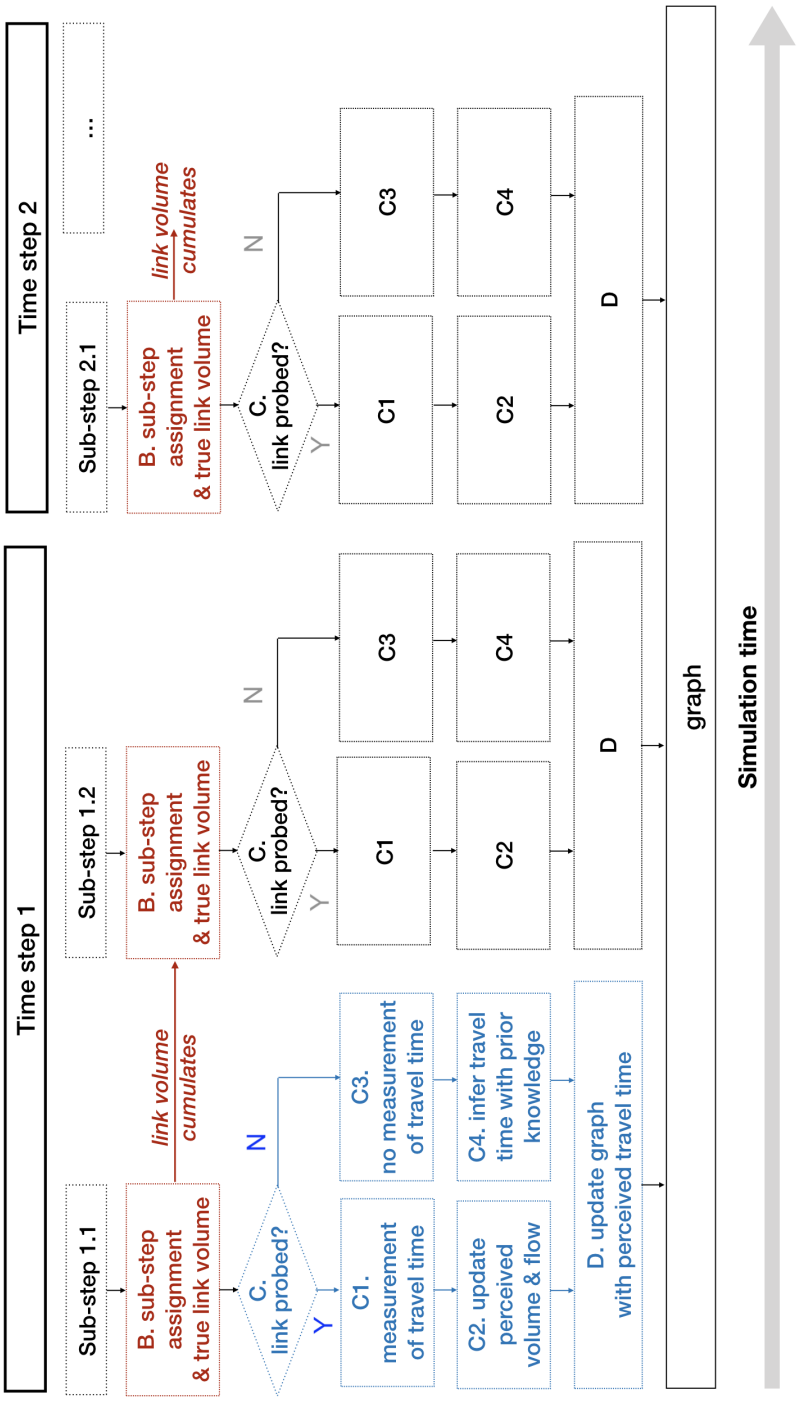


Fig. 4.3 Simulation framework considering partial knowledge of traffic conditions from probe data.

where $t_{t,l}$ is the final reported link-level travel time for link l according to all individual probe measurements $t_{t,l,p}$. Subscript t, l and p means travel time in traffic conditions (as opposed to the free flow condition), link index and probe index, respectively. N_p is the number of all the probes that go pass link l. Link-level travel time measured by individual probe is a stochastic variable that follows a distribution, whose expectation $\mu_{t,l}$ is the true average travel time given by the BPR function as a function of the true link flow (Step B). The variance term of the of the distribution $\sigma^2_{t,l}$ is related to the variability of individual driving patterns. Both the distribution and the variance term used in the simulation will be introduced in more detail in the simulation set-up section. The traffic managers receive the individual $t_{t,l,p}$ and aggregate it by link to obtain $t_{t,l}$, which will be used by the next batch of travellers for route planning.

Another process is needed so as to calculate and update the traffic managers' knowledge of the link-level congestion level, i.e., the flow rate and volume. Such knowledge is not required for the updating of graph weights in this simulation sub-step, but is crucial for the estimation of link-level travel times in future sub-steps when no probe measurement is available. The BPR curve (Equation 3.3 in Section 3.5.2) provides a monotonically increasing relationship between the flow rate and travel time, so the link-level flow rate (numbers of vehicles passed in a unit time, e.g., one hour) can be uniquely determined and promptly updated based on the measured travel time. Closely related to the flow rate is the traffic volume, which is the vehicle count not normalised by time and represents the total number of vehicles that have passed a link since the beginning of the current simulation time step. The link-level volume can be calculated by scaling the link-level flow rate through inversely using the relationships in Equation 3.2a and 3.2b in Section 3.5.2. This assumes global knowledge by the simulator of the overall progress, replacing the time elapsed (in a true dynamic simulation) by the proportions of trips assigned within a time step (in the static/pseudo-dynamic simulation with sub-stepping).

C3 & C4. Updating travel time and traffic volume for unprobed links

If a link is not probed in a particular sub-step, the traffic managers can still update the estimated link-level travel time based on some prior knowledge, though with a different procedure than for the probed links. This process starts with estimating the current congestion level, e.g., link-level traffic volume. As no probe has passed the link, the traffic managers are thus observing no increase in the vehicle counts or traffic volume. So the traffic volume of an unprobed link is not updated. If a link has never been probed since the initiation of the time step, the perceived traffic volume of that link always equals to zero.

After obtaining (retaining) the link-level traffic volume for the current sub-step, the hourly flow rate can be updated according to 3.2b in Section 3.5.2. Note that in 3.2b, there are two

other variables being used, namely the *demand assigned* and *hourly demand*. This is not necessarily assuming the traffic managers have a knowledge of the travel demand a priori, but rather to use the ratio between these two volumes to represent the progress of the simulation in replacement of time. Currently, it is unable to achieve the truly dynamic mesoscopic traffic simulation due to the large computational efforts (the increased frequencies of shortest path finding; the iterations to assure the consistency between the distance a trip takes and the time elapsed in the simulation). As a result, the minimum time unit used is one hour. Trips within each hour are assigned incrementally without the notion of time. In the future, if the duration of each time step is small enough and no sub-stepping is involved (i.e., truly dynamic traffic assignment), the ratio between *demand assigned* and *hourly demand* can well be replaced by $time\ step\ duration/1\ hour$.

Based on the estimated link-level flow rate, the estimated link-level travel time can be obtained by applying the BPR function (Equation 3.3 in Section 3.5.2). Overall, in the unprobed case, the estimated link-level volume is less than or equal to the true volume, thus the hourly flow rate and link-level travel time both tend to be underestimated as well.

4.5 Simulation Set-up

Traffic simulation experiments in this chapter are designed to answer three questions that are not fully addressed in Chapter 3. Specifically, they are:

1. Influence of the probe penetration rate on traffic simulation outcomes (traffic distribution and vehicle hours travelled, VHT). This process is the same as seeking the minimum probe penetration rate required to obtain the information and achieve an accepted level of efficiency of the traffic system.

2. Influence of the probe information variability. As the mesoscopic traffic simulation relies heavily on the BPR volume-delay curve, naturally it does not embed variabilities in vehicle-to-vehicle driving pattern. As long as the link flow and other link attributes are the same, the BPR curve will lead to a fixed travel time. This effectively means that all vehicles travelling through a particular link in the same time step will experience the same time duration to go through a road link. While by considering probe information in this chapter, it should also be included in the simulation that a larger number of probe vehicles will lead to more reliable travel time information. As a result, variabilities in travel time experienced by individuals are also investigated within the scope of this chapter.

3. The influence of the random trip assignment order and stochastic individual probe variability. In each time step, trips are randomly dispatched in 20 sub-steps. Even though the number of 20 increments is much larger than in similar studies [154] and is not expected to cause big variations in the network-level distributions of traffic, this postulation will be tested in this chapter. In addition, the variabilities in the probe data may also lead to variations in the outcome variables. The sensitivities of the modelling results to the stochastic process involved are evaluated by repeating the same simulation under the same parametrisation for several times, and by visualising the dispersions in the outcomes.

To reduce the computational effort, only one day of traffic simulation, as opposed to a whole week in Chapter 3, is used in this chapter for the probe information study. Friday (3 AM till 2 AM on Saturday) is chosen due to its representativeness of a full range of traffic conditions. It has the near free flow demand at 3 AM while the traffic demand at 6 PM is the highest among all 7×24 hours in a whole week. Again, as in Chapter 3, the 24 hour simulation starts from 3 AM rather than the midnight as the traffic at 3 AM is the lowest, thus justifying the initialisation of the weight of road network graph with the free flow travel time of each link.

Probe penetration rates

Six probe penetration rates are studied. They include 0% (no information), 0.1%, 0.5%, 1%, 10% and 100%. Specifically, the 100% case corresponds to the set-up in Chapter 3. At each simulation time step, a proportion of vehicles equalling to the penetration rates are randomly sampled out to act as probes. Table 4.1 lists the hourly and total vehicle counts and probe counts according to the travel demand on a typical Friday. The number of trips is the lowest at 3 AM (nearly 10,000 trips), which gradually reaches the morning peak at 8 AM (nearly 100,000 trips). The travel demand decreases through the mid-day and reaches the evening peak at 6 PM (more than 170,000 trips). The rest of the columns in Table 4.1 show the number of probe vehicles in each hourly time step according to the probe penetration rate scenarios. Depending on the random seeds used, the actual probe counts may vary a little.

Probe information variability

The BPR curve assumes a deterministic relationship between link volume and traversal time. However, this is hardly realistic given the diverse driving patterns of individuals, the unpredictable delay due to traffic lights and so on. Even though it is assumed that the BPR curve will give a mean estimation of the travel time at a particular traffic congestion state

Table 4.1 Hourly and total vehicle counts and probe counts used in the simulation

Hour	Journey counts	Probe penetration rate					
		0%	0.1%	0.5%	1%	10%	100%
3 AM	9,458	0	10	46	88	937	9,458
4 AM	9,375	0	7	35	83	893	9,375
5 AM	15,757	0	12	63	143	1,552	15,757
6 AM	30,068	0	25	148	301	2,943	30,068
7 AM	76,981	0	90	395	758	7,705	76,981
8 AM	99,730	0	113	528	1005	9,885	99,730
9 AM	79,698	0	99	415	786	7,890	79,698
10 AM	58,758	0	77	295	548	5,930	58,758
11 AM	54,350	0	57	252	498	5,406	54,350
12 PM	56,260	0	49	255	528	5,559	56,260
1 PM	56,831	0	68	261	518	5,681	56,831
2 PM	60,135	0	73	318	617	6,090	60,135
3 PM	68,510	0	65	325	680	6,816	68,510
4 PM	83,006	0	89	419	847	8,347	83,006
5 PM	139,919	0	144	745	1,448	13,860	139,919
6 PM	172,943	0	161	836	1,707	17,281	172,943
7 PM	148,462	0	152	742	1,522	14,771	148,462
8 PM	127,063	0	118	602	1,256	12,720	127,063
9 PM	119,190	0	108	608	1,187	11,937	119,190
10 PM	121,793	0	111	619	1,190	12,149	121,793
11 PM	116,844	0	113	580	1,180	11,706	116,844
12 AM	101,842	0	112	497	981	10,327	101,842
1 AM (Saturday)	79,219	0	88	427	847	7,917	79,219
2 AM (Saturday)	44,997	0	28	200	427	4,465	44,887
Total	1,931,079	0	1,969	9,611	19,145	192,767	1,931,079

(given by the flow-to-capacity ratio), it is likely that individual probe vehicles will experience and report a different travel time.

Travel time variability is made of several components, including (1) day-to-day variation (caused by, e.g., incidents, road work closures or change in weather conditions), (2) within-day variation (caused by different congestion levels) and (3) vehicle-to-vehicle variation (caused by heterogeneity in driving patterns, delays at intersections, etc.) [127, 168]. In this study, it is the vehicle-to-vehicle variation component that is of interest, as the perceived link-level travel time is calculated based on the probe travel time from the previous time step, rather than the previous day or over a course of several days. In other words, the day-to-day variation is simply not considered as the traffic simulation is already assumed to be based on typical weekday or weekend travel demand, while the within-day variation will be accounted for by the varying hourly demand inputs. Many previous literatures studying the reliability of travel time target at the first two components, thus are not particularly relevant to this study [51, 58, 223].

As for the vehicle-to-vehicle travel time variability, the most recent and systematic studies are found in Kim and Mahmassani [127], Mahmassani et al. [148]. Vehicle-to-vehicle travel time variability refers to the variations in travel time experienced between drivers who depart within the same time slot, thus excluding exogenous factors such as peak vs non-peak hour variations (within-day variability) or the delays due to accidents (day-to-day variability) [127]. The vehicle-to-vehicle variability in [127] is quantified based on the travel delay per unit distance x:

$$x_{t,l,p} = \frac{t_{t,l,p} - t_{f,l}}{link\_length} \tag{4.3}$$

where $t_{f,l}$ is the free flow travel time on link l and $t_{t,l,p}$, as in Equation 4.2, is the experienced travel time of driver/probe p on link l. Correspondingly, $x_{t,l,p}$ is the travel delay per unit distance experienced by driver/probe p on link l. Kim and Mahmassani [128] proposed a multiplicative error structure for $x_{t,l,p}$ with the variability term γ following a *Gamma* distribution (Equation 4.5). Thus $x_{t,l,p}$ also follows a *Gamma* distribution with shape parameter π and scale parameter $\mu_{xt,l}/\pi$.

$$\begin{aligned} x_{t,l,p} &= \mu_{xt,l}\gamma \\ \gamma &\sim Gamma(\pi, 1/\pi) \end{aligned} \tag{4.4}$$

$$x_{t,l,p} \sim Gamma(\pi, \mu_{xt,l}/\pi) \tag{4.5}$$

$$\text{With Mean: } \pi \times \mu_{xt,l}/\pi = \mu_{xt,l}$$

$$\text{Standard deviation: } \sqrt{\pi} \times (\mu_{xt,l}/\pi) = \mu_{xt,l}/\sqrt{\pi}$$

$$\text{COV: } \frac{Standard\ deviation}{Mean} = 1/\sqrt{\pi}$$

$\mu_{xt,l}$ is the mean of the unit travel delay $x_{t,l,p}$ and can be obtained from the BPR function. Besides, in Mahmassani et al. [148], simulations are conducted based on three real networks and it is found that the standard deviation ($\mu_{xt,l}/\sqrt{\pi}$) can be expressed as a linear relationship of $\mu_{xt,l}$, i.e, the coefficient of variation (COV) is a fixed value. Results of these simulations and real world validation based on GPS trajectories suggest that the COV of $1/\sqrt{\pi}$ can be as low as 0.4668 or as high as 1.0358.

Based on the above, three COVs are tested in the simulation of probe data variability: 0.5, 1.0 and 2.0. An intuition of the COV is as the following: Suppose the free flow travel time on link L is 60 seconds. Based on the current congestion level, the BPR curve predicts an average delay of 20 seconds, making the expectation of the traversal time 80 seconds. Given a COV of 1.0 to be tested (making $\pi = 1$ in the *Gamma* distribution), the standard deviation of an individual traversing link L is $20 \times 1 = 20$ seconds, while the distribution of the individual experienced travel time $\sim Gamma(1, 80)$. In case there are multiple probes, the average of their experienced travel times is used for the final reporting. The average of N *i.i.d Gamma*-distributed variables still follows a *Gamma* distribution, with the same expectation $\mu_{t,l}$ but a smaller standard deviation.

$$t_{t,l} \sim Gamma(N \cdot \pi, \frac{\mu_{t,l}}{N \cdot \pi}) \tag{4.6}$$

Figure 4.4(a) shows the probability density distribution plots of the variability term γ under different COV value assumptions. To visualise the variabilities in individual vehicle travel times in a more straightforward manner, Figure 4.4(b) shows a sample of 6000 points of $x_{t,l,p}$, given different $\mu_{xt,l}$ and COV values. It can be seen that when the COV value is small (e.g., COV=0.5), γ is more concentrated to 1 (Figure 4.4(a)) and the vehicle-to-vehicle speed variabilities are small (Figure 4.4(b)). For a relatively big COV (e.g., COV=2.0), γ is more spread out, especially to areas with $\gamma \geq 1$ (Figure 4.4(a)). As γ is the ratio between $x_{t,l,p}$ and $\mu_{xt,l}$, the unit delays of some vehicles can become very large (Figure 4.4(b)).

In summary of the simulation set-up, traffic is simulated based on a typical Friday demand. Six levels of probe penetration rates (0%, 0.1%, 0.5%, 1%, 10% and 100%) and four levels

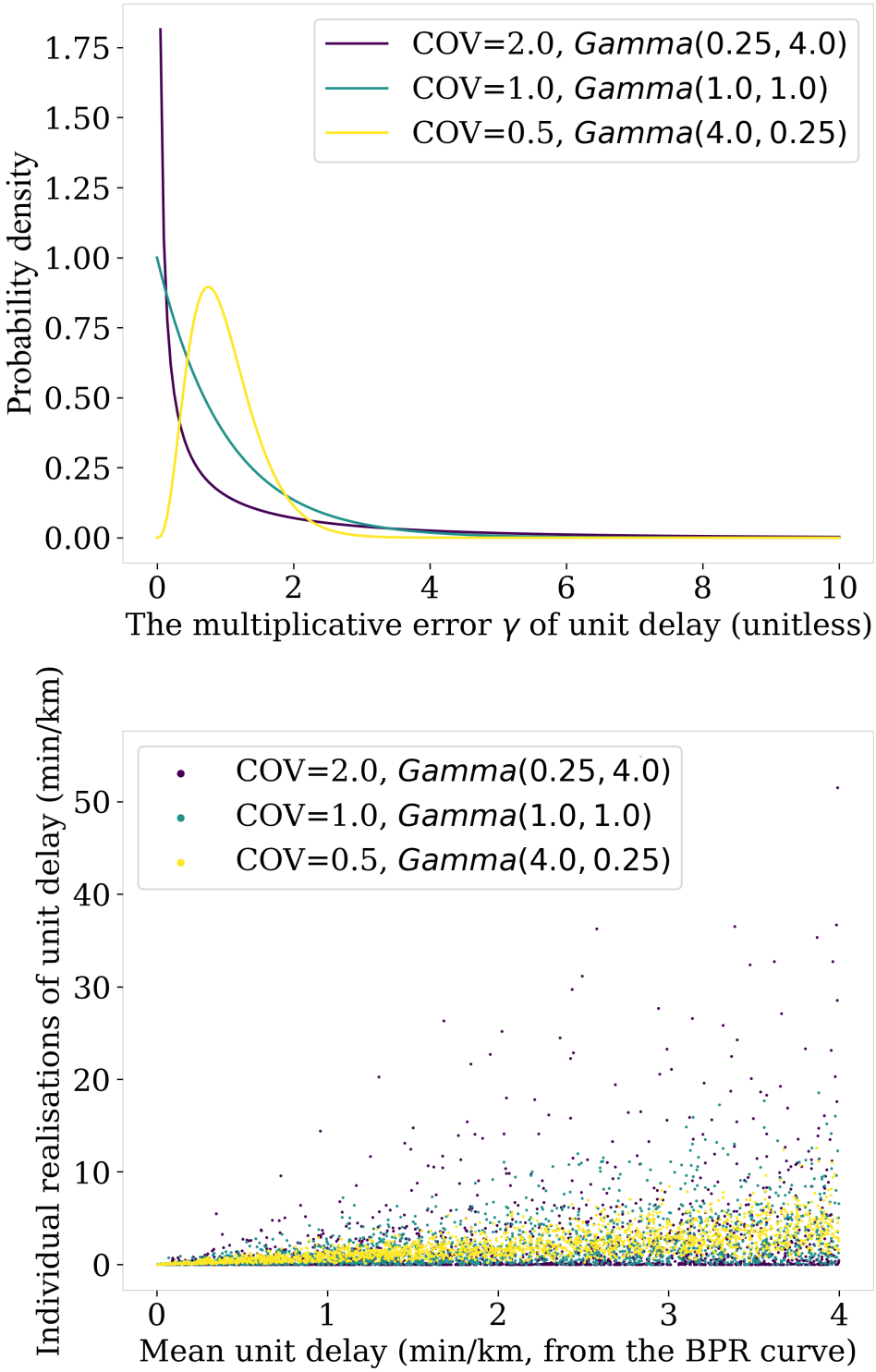


Fig. 4.4 Distributions of individual vehicle's unit delay. (a) Probability density plots of the variability term γ, which is the ratio between the individual vehicle's unit delay $x_{t,l,p}$ and the average unit delay to be expected on a link $\mu_{xt,l}$. (b) Scatter plots of stochastic realisations of $x_{t,l,p}$ given $\mu_{xt,l}$ and different COVs. COVs and γ are related as shown in (a).

of probe information variations (COVs: 0, 0.5, 1.0 or 2.0) are specified. In addition, each combination of parameters is repeatedly simulated for 10 times with different random seeds to quantify the variations in results due to the random and stochastic processes involved in the simulation.

4.6 Results

As stated in Section 4.1, the primary purpose to include the analysis of this chapter is mainly to evaluate the suitability of the traffic simulation assumptions for later applications, including the assumption of "perfect live traffic information" for making route choice and the variability in the results due to the random trip assignment orders by sub-step. Depending on the results of this chapter, the assumptions may need to be modified for the subsequent analysis in this thesis. While a secondary aim of this chapter is to use the mesoscopic traffic simulation as a standalone tool (not coupled with other infrastructure models) for traffic analysis applications. Particularly in this section, the results will be interpreted from the traffic analysis perspective, as it is easier to present the numbers with a real context. The evaluation of the primary purpose, i.e., whether the model assumptions are suitable for subsequent analysis, will be given in the discussion section of this chapter (Section 4.7).

4.6.1 Probe penetration rate: its effects and the minimum coverage requirement

To demonstrate the effects of various probe penetration rates on the traffic system efficiency, the VHT are plotted together with a fixed probe data variability. The VHT is a typical measure of system efficiency and Figure 4.5 displays the hourly total VHT throughout the day when the probe data variability is 0. Specifically, the horizontal axis is the departure hour (24 hours, from 3 AM on Friday to 2 AM on Saturday). The vertical axis is the total VHT of all trips departed in one specific hour. The six coloured series represent the six levels of probe penetration rates from 0% to 100%. The solid lines indicate the median VHT based on 10 repetitions of the simulations. The box and whiskers show the variabilities of the first and third quantiles (box) and plus/minus one interquartile range. For data located outside of this range, they are plotted as outlier points.

By observing Figure 4.5, clear hourly variations of VHT can be seen, including the hump at the morning peak hours (7 - 9 AM) and a even higher one at the evening peak hours (5 - 8 PM). Figure 4.6 displays the same information, but the total VHT has been divided by the hourly trip counts to obtain the average trip travel time. It can be seen that for trips

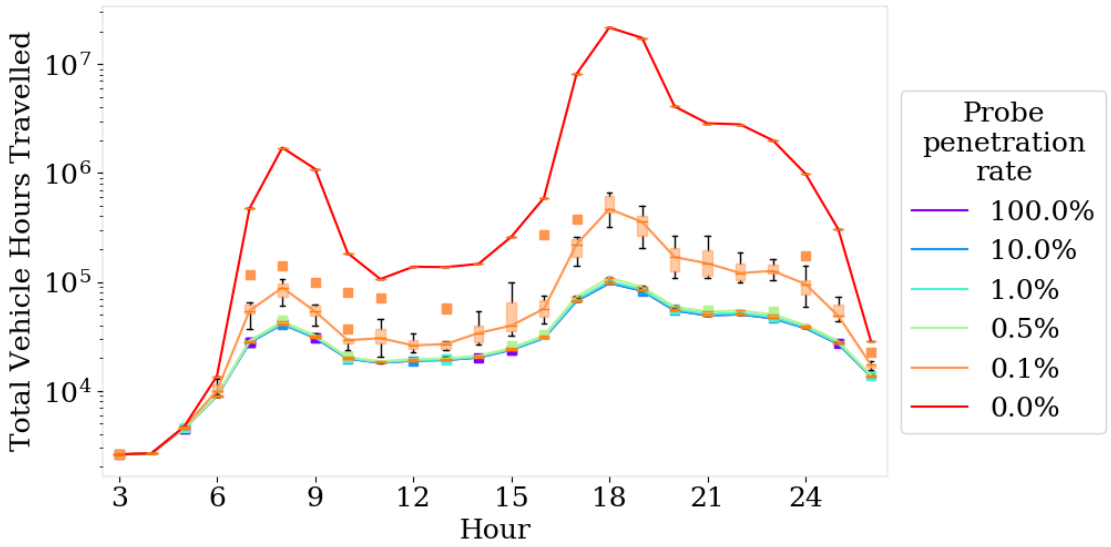


Fig. 4.5 Total VHT by hour on a typical Friday, given different probe ratios yet no variation in probe information. Specifically, no error bar is drawn for the 0% probe penetration rate case (the red series). This is because the outcomes are deterministic and all the vehicles always make same route choices based on the free flow conditions.

happening during the peak hours, they experience more delays compared to those that are taken in off-peak hours.

A second observation of Figure 4.5 and 4.6 is that a minimum of 0.5% of probe vehicles are required to ensure the traffic efficiency. In fact, the traffic efficiency does not change much when the probe penetration rate is above 0.5%, indicating a depreciated value of information once the quantity has passed a threshold. However, when the probe penetration rate is 0.1% or 0% (the "no information" case), the VHT or the average travel time becomes several orders higher. This can be investigated together with Table 4.1: When the probe penetration rate is 0.1%, there are only less than 100 vehicles in the system in each time step. This turns out to be insufficient in providing the necessary live traffic condition updates. As a result, many travellers are not able to take the optimum path. This is shown by the traffic distribution maps in Figure 4.7. As seen from the maps, when the probe ratio is close to 0 (Figure 4.7 (a) and (b)), the traffic on highways are significantly higher than in other cases. According to the BPR curve, when the traffic volume doubles on a link, the travel time delay is penalised to the power of four, leading to an excess of VHT (travel time $\times$ traffic volume) to be more than 30 times higher. In fact, with an average trip travel time well above 60 minutes for most hours of the day, the outcomes with probe penetration rates of 0.1% and 0% are not realistic given the fact that SF roughly measures 11 km by 11 km. With a probe penetration

rate higher than 0.5%, the traffic efficiency can be assured, thus the traffic distributions are approximately optimum and hardly distinguishable in Figure 4.7 (c)-(f).

Another noticeable observation based on Figure 4.5 and 4.6 is that the variations in the results are not significant except for the case with 0.1% probe penetration rate. The randomness/stochasticity in the simulation comes from three sources: (1) the random selection of probes from all vehicles; (2) the random vehicle dispatch order within one time step (traffic assignment order according to the sub-steps) and (3) the stochastic probe data variability. As (2) is more related to the general mesoscopic modelling assumption, it will be discussed in the discussion section (Section 4.7). While in this set of simulations, no probe data variability is assumed yet, the variability also does not come from (3). The effect of (3) will be described subsequently in the next subsection. Consequently, this leaves the only explanation of the variations, most noticeably around the case with 0.1% probe penetration rate, to (1). Due to the low numbers of probes in this case, only the real-time traffic information of a small fraction of roads is collected. Depending on which travellers are labelled as the probes, information of completely different parts of the network may be revealed. In contrast, when the probe penetration rate is higher, the coverage of the whole traffic dynamics also becomes better. And from this analysis, it is found that 0.5% is sufficient to provide this coverage. Lastly, when the probe ratio is 0%, all travellers will choose the same routes that they use in the free flow conditions, thus their travel times will not change.

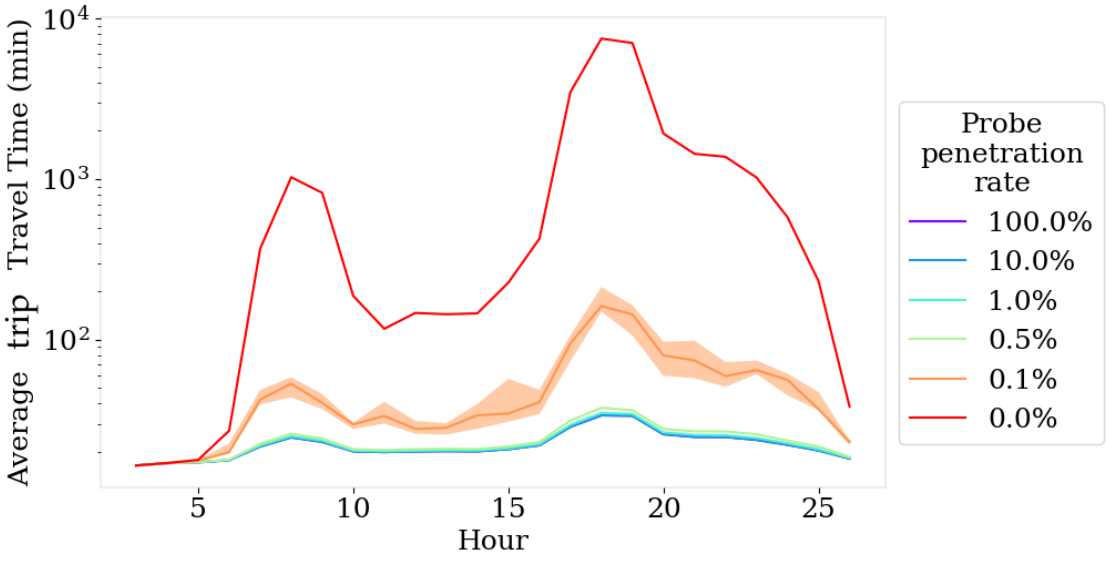


Fig. 4.6 Average trip travel time (minutes) by hour on a typical Friday, given different probe ratios yet no variation in probe information. As in Figure 4.5, no error bar is drawn for the 0% probe penetration rate case (the red series). This is because the outcomes are deterministic and all the vehicles always make same route choices based on the free flow conditions.

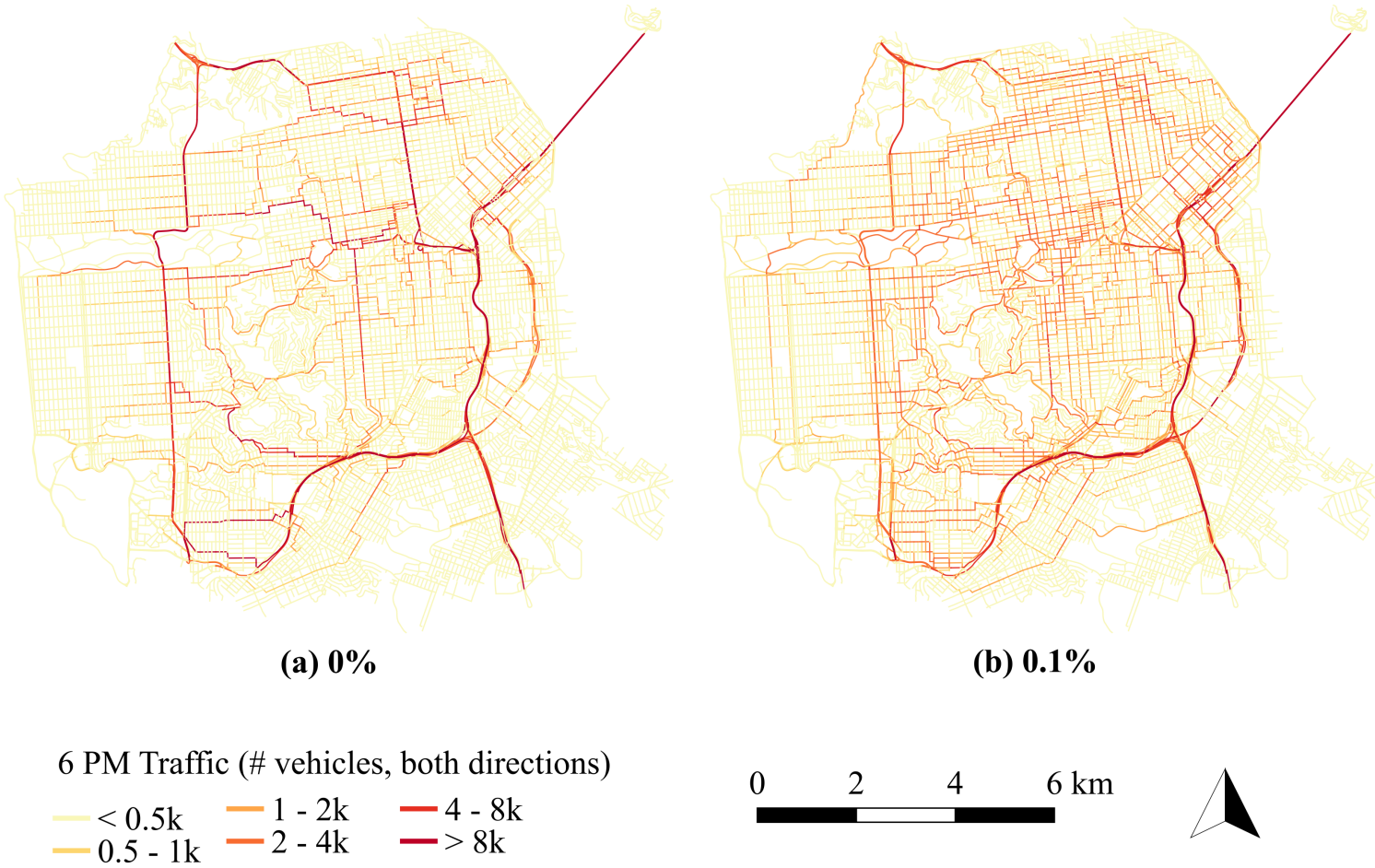


Fig. 4.7 Traffic volume at 6 PM given different probe ratios but no variation in probe information (to be continued in the next page). (a) Probe ratio is 0 (no probe information). (b) Probe ratio is 0.1%. Traffic is more distributed in (b).

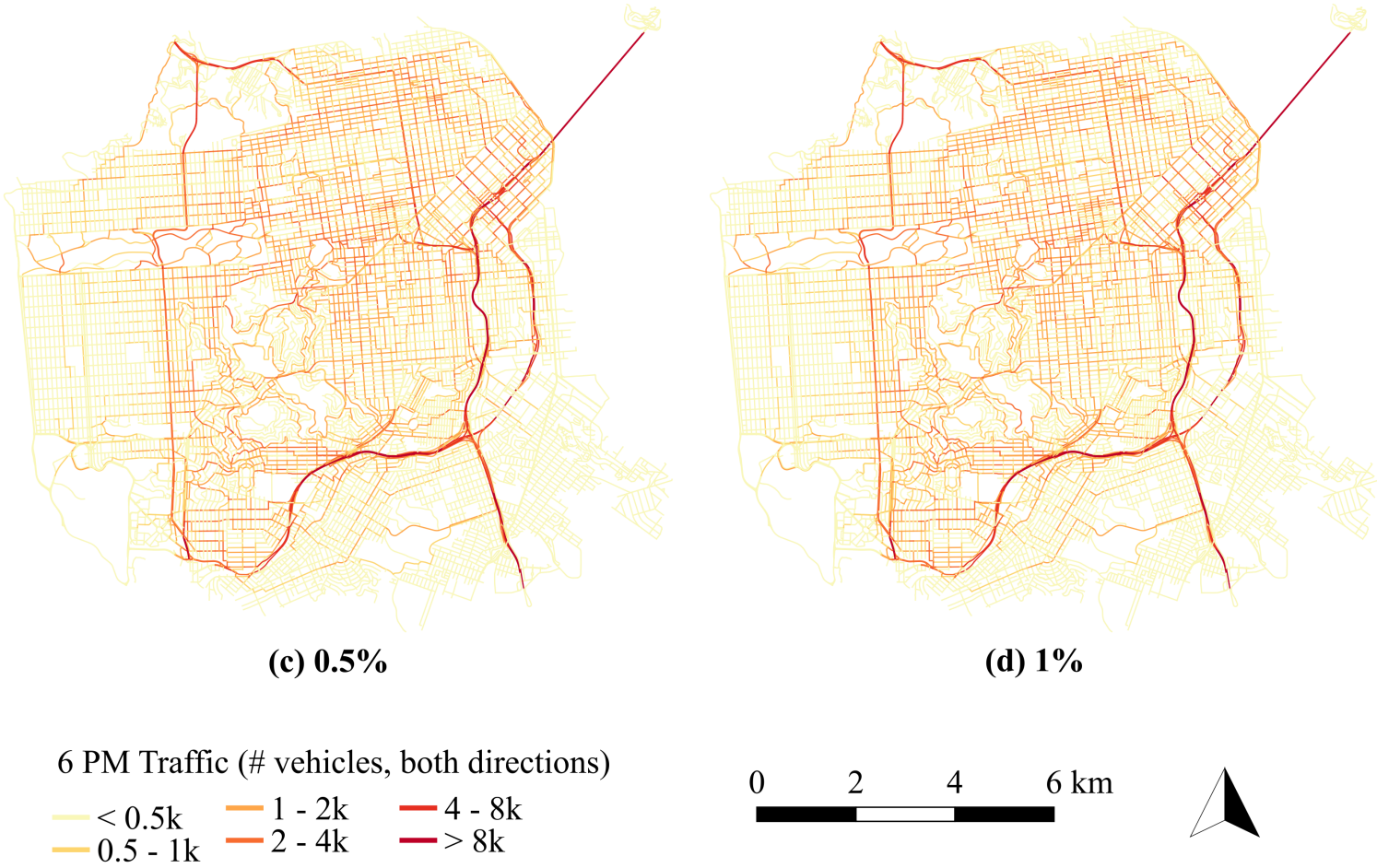


Fig. 4.7 Traffic volume at 6 PM given different probe ratios but no variation in probe information (cont. and to be continued in the next page). (c) Probe ratio is 0.5%. (d) Probe ratio is 1%. As there is already sufficient information to ensure system efficiency in both cases, the traffic distributions in these two maps are very similar.

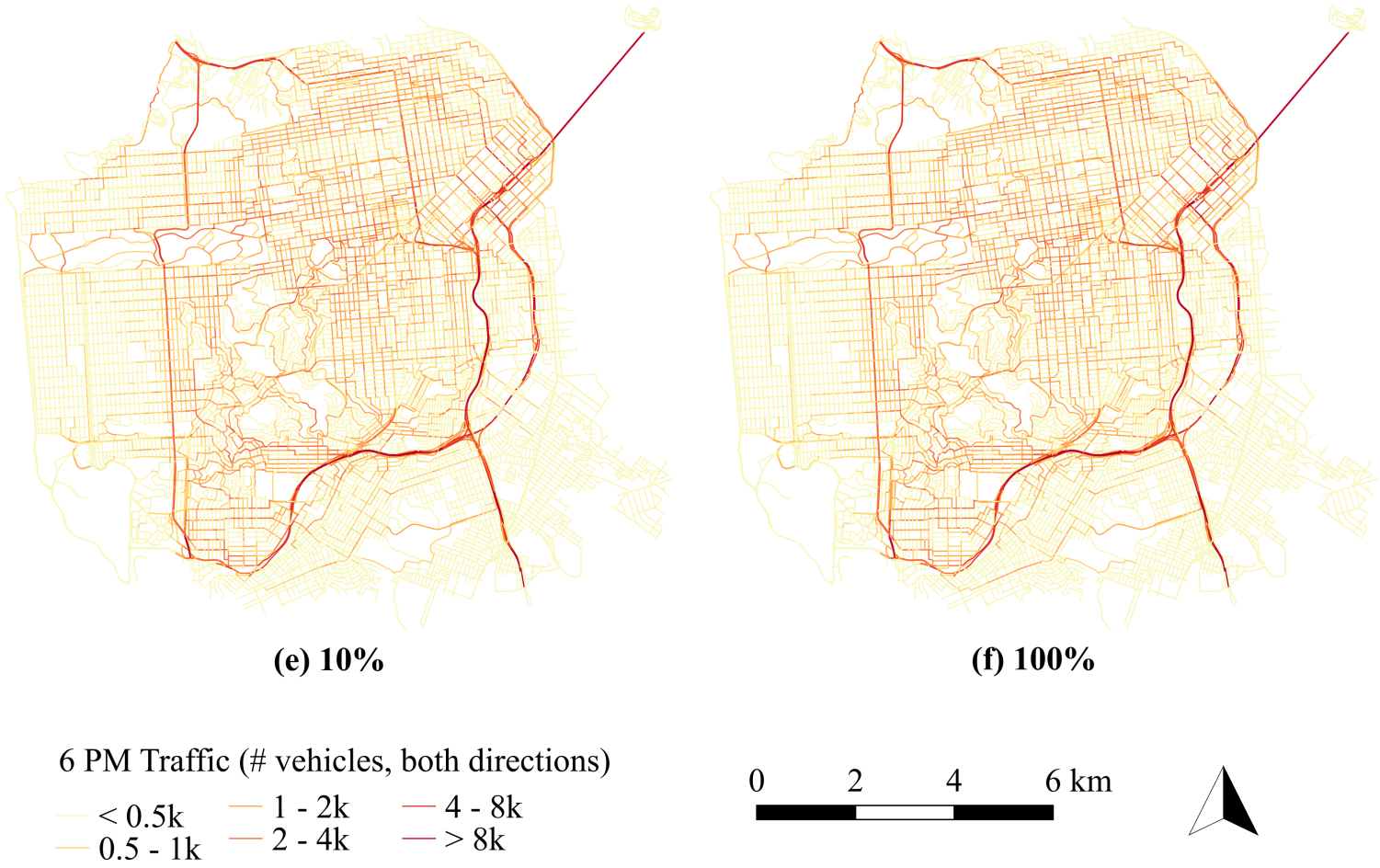


Fig. 4.7 Traffic volume at 6 PM given different probe ratios but no variation in probe information (cont.). (e) Probe ratio is 10%. (f) Probe ratio is 100% (full probe information - corresponding to the case in Chapter 3). As there is already sufficient information to ensure system efficiency in both cases, the traffic distributions in these two maps are very similar.

The information in Figure 4.7 is plotted in parallel coordinates plot in Figure 4.8. The road links in the 2D maps are reduced to points on 1D vertical axes, with the position arranged according to the hourly traffic flow. The lines between two adjacent axes connect two points corresponding to the same road link. As a result, Figure 4.8 shows the changes of the traffic distribution across six different probe ratios. The red lines highlight those links whose hourly volumes at 6 PM on a typical Friday are among the top 0.1% of all road links. The blue lines represent links with hourly traffic volumes at 6 PM above the 99th percentile but below the top 0.1%. Cyan links are those with hourly traffic volumes at 6 PM between the 90th and 99th percentile, while black lines denote the reset of the links. It can be seen that for the first four parallel coordinates (probe penetration rate above 0.5%), the coloured lines are almost all horizontally aligned, indicating no change in the hourly flow on any of these links despite a varying level of probe penetration rates. In other words, decreasing the probe penetration rate from 100% to as low as 0.5% does not cause noticeable change in the traffic distributions. However, when the probe ratio is 0.1%, a small dispersion trend can be seen in the lines, which accentuates into a wider dispersion in the last parallel coordinate axis. In these last two vertical axes, the red and blue lines tend to go higher, while the cyan lines grow lower. Figure 4.9 shows the result by masking out the red, blue and cyan lines and it can be seen that the black lines are also dipping down toward the last two parallel coordinates. This echoes the observation of Figure 4.7. The red and blue lines in the parallel coordinate plot represent the top 1% of the road links in terms of traffic flow in the perfect information case (probe penetration rate being 100%). These are usually highways and important arterial roads. For the trend of the red and blue lines to go up as the probe penetration rate decreases, it means that these highways and arterial roads are carrying even heavier traffic, just as displayed in Figure 4.7 (a) and (b). Similarly, the cyan and black lines represent those links that are relatively less important in the road network. A low probe penetration rate below 0.1% sees a decrease of traffic volume in this part of the road network. Since the traffic volumes are plotted in the *log* scale, the absolute variations in the last two parallel coordinates are actually more profound in the absolute term.

This observation from the parallel coordinate plot is logical given that without probe information, travellers will choose routes based on the free flow travel time, which usually favours the highways as they have a higher speed limit. Figure 4.8 and 4.9 are able to visualise this effect in a more disaggregated form. As previously found in Figure 4.5, 0.5% is the critical probe penetration rate in this study, below which the traffic distribution is considerably different from the perfect information case and the efficiency of the transport network is compromised.

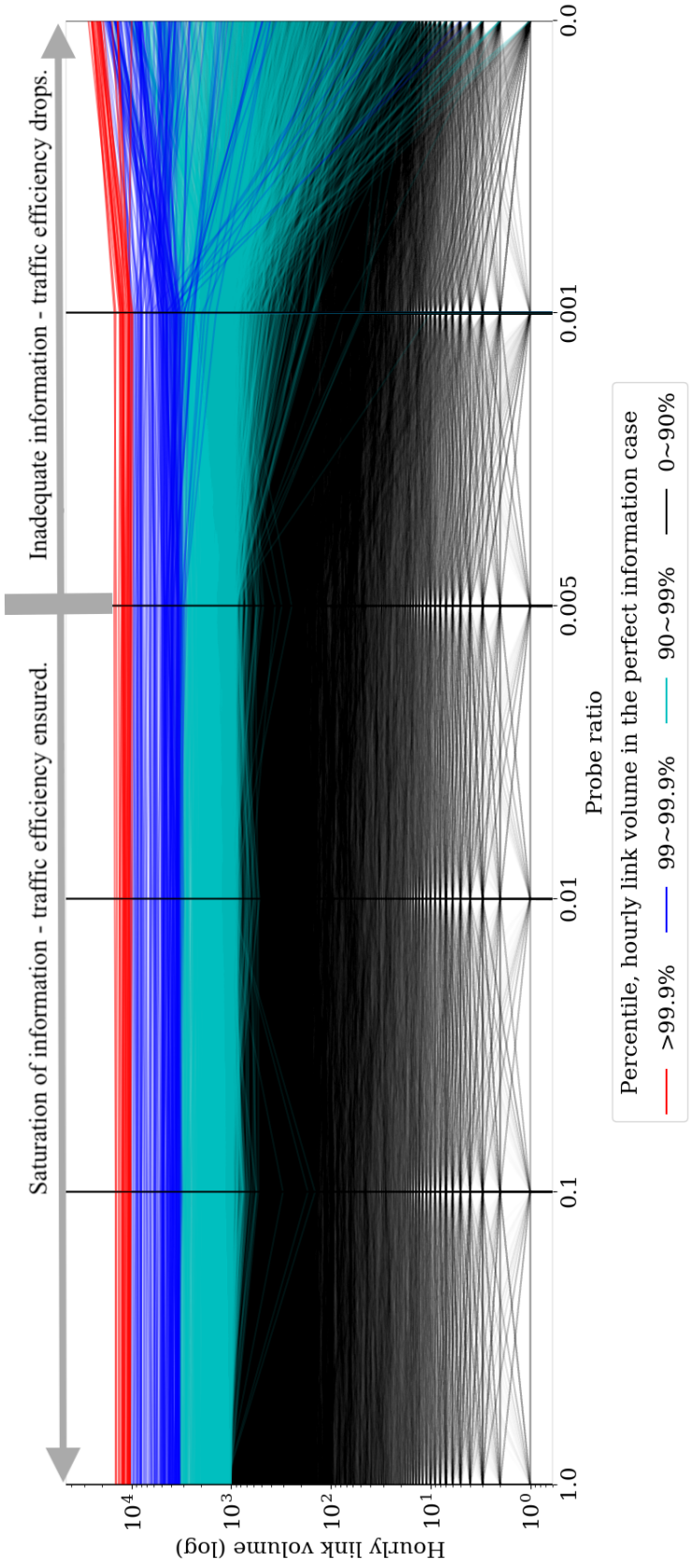


Fig. 4.8 Parallel coordinate plot of 6 PM link usage variation across different probe ratios, coloured by link-volume percentiles when probe ratio is 100%.

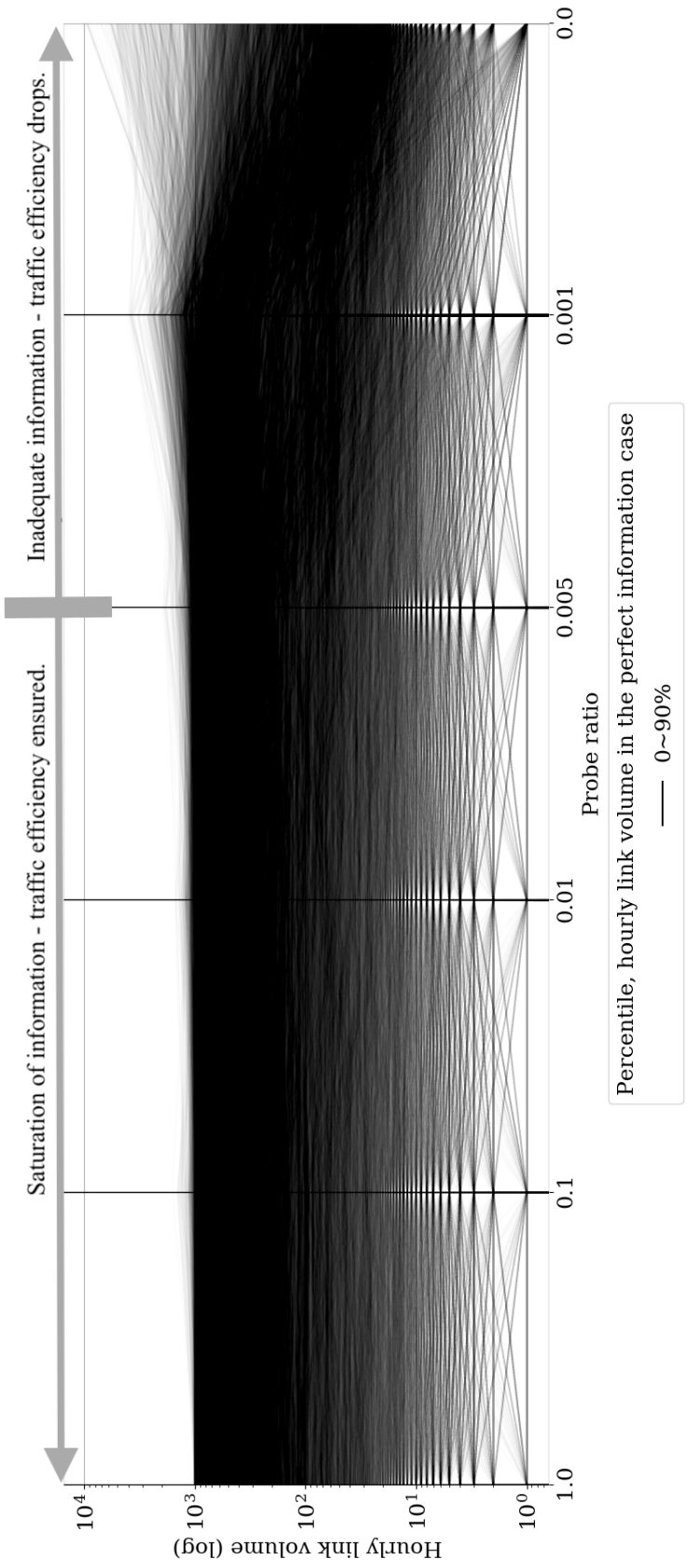


Fig. 4.9 Parallel coordinate plot of 6 PM link usage variation across different probe ratios, only showing links with hourly volume percentiles below 90%.

4.6.2 Impacts of probe information variability

The success of traffic information system relies on the reliability of the data it collects and disseminates. When the variability in the probe data measurements are significant, it may provide a misleading knowledge about the traffic conditions. In this subsection, results of the impacts of probe data variability, measured by the changes in travel time, are presented. Since the previous subsection has demonstrated that the outcomes based on a probe ratio of 0.1% or 0% are far from being realistic, only results from the cases with probe ratios $\leq 0.5\%$ are presented in this subsection.

The changes in the average trip-level travel time are shown in Figure 4.10. The probe ratio is fixed at 0.5%, the minimum acceptable requirement for this study, while four probe data variabilities are plotted in the data series. The probe data variability is specified by the COV of the unit length delay (Equation 4.3 and 4.6). The series labelled with a COV of 0 (the purple line) corresponds to the case when no individual variability from the mean link-level delay exists in the 0.5% of probe vehicles. All the data series are close to each other and the largest separation of the data series occur at 6 PM.

Taking a slice of the data series shown in Figure 4.10 at the 6 PM scenario and combining these data points with those cases of different probe penetration rates, Figure 4.11 visualises the outcomes of the average trip travel time at 6 PM of a typical Friday given different combinations of probe penetration rates and information variabilities. In the best case, when the probe penetration rate is 100% and no individual variability exists (the lower right corner of the plot), the average journey time for trips occurring at the 6 PM evening peak is 33.8 minutes. As the probe penetration rate decreases to 0.5%, the average journey duration is around 10% longer at 37.5 minutes (purple series). While with a large probe information variability, the system is becoming even less efficient and the average trip duration can be as high as 40.3 minutes when the probe COV becomes 2 (19% longer than the best case). In real life, the variations in the probe information may come from driving behaviours, delays at traffic signals, interactions with buses, pedestrians or even the false data collected from a pedestrian or motorist with GPS enabled mobile devices. All these unpredictable delays may generate misleading information as for whether a road link is really congested. This highlights the need of a higher probe penetration rate or other measures to filter out instantaneous anomalies.

The box and whisker plots in Figure 4.10 as well as in Figure 4.11 again shows the range of the first and third quantiles based on 10 repetitions of the same parameter set. It can be seen that small fluctuations in the outcomes due to the stochastic process exist in all cases and become wider at a low probe penetration rate (e.g., 0.5%).

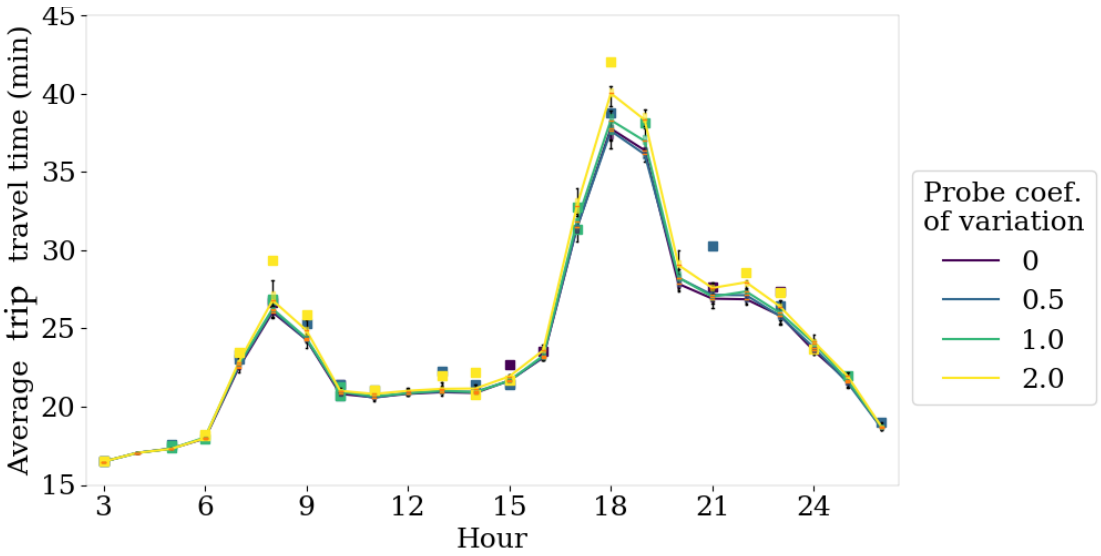


Fig. 4.10 Average trip travel time (minutes) by hour on a typical Friday, given fixed probe ratio (0.5%) and different levels of variations probe information.

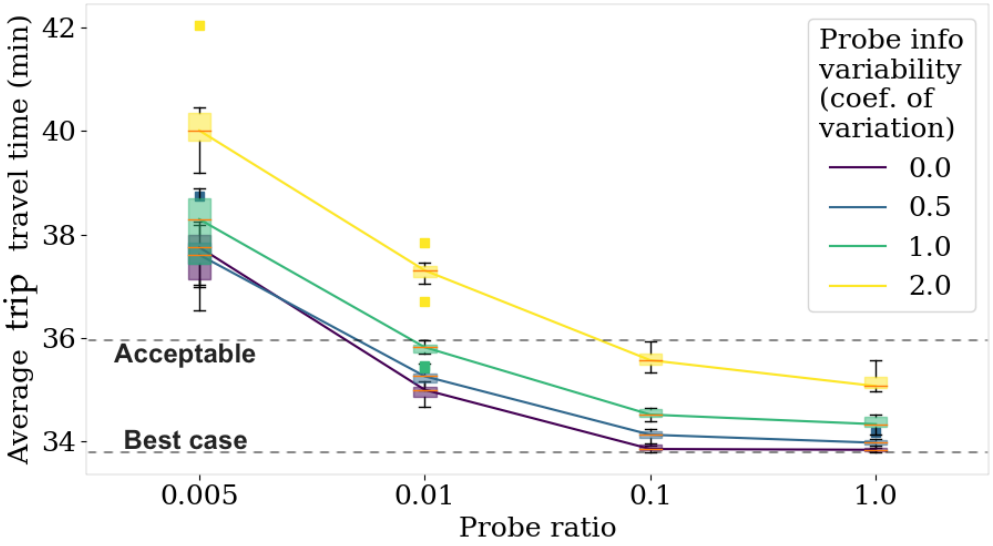


Fig. 4.11 Average trip travel time (minutes) at 6 PM on a typical Friday, given different combinations of probe ratio (0.5%, 1%, 10%, 100%) and variations in probe information (coef. of variation of unit length delay = 0, 0.5, 1.0, 2.0).

4.7 Discussions

The plausibility of the assumption of perfect information

The assumption of access or awareness of the best route is the basic assumption adopted in many traffic models, since the equilibrium models devised in the early days. As shown

in this chapter, the analysis process can be greatly simplified without considering the data availability issues. The minimum acceptable probe ratio found is 0.5% when no individual variability is presented. This is lower than the threshold reported elsewhere (5-7%) [87, 226]. This discrepancy is mainly because the time step used in this analysis (one-hour) is quite long for this type of analysis, and assuming no sub-hour dynamics (e.g., change in link-level travel time minute-by-minute) would lead to a lower requirement of the probe ratio. But in either case, the minimum required probe ratio is a low number compared with the smart phone ownership rate in the study area.

When adding in the effect of probe data variability, the requirement of the probe ratio is slightly higher. For example, in the best case the average trip travel time is around 34 minutes. An extra 1 or 2 minutes longer than the best case may deem acceptable on an individual level as a similar amount of time might as well be lost on parking, etc. As shown by the dashed lines in Figure 4.11, this requirement sets the minimum desired probe ratio higher than 0.5% in most cases, to about a few percent. Still, this requirement is considered to have been met due to the high numbers of smart phone users in an urban area. As a result, the assumption of perfect information is realistic and testable in this study.

The impact of random trip assignment order

As shown by the model structure of the mesoscopic simulation, trips in each hourly time step is randomly divided into 20 batches (sub-steps). The traffic assignment of these 20 sub-steps are carried out sequentially to achieve an equilibrium-like traffic pattern than the all-or-nothing static assignment process (without sub-stepping, all travellers in a time step are assigned in one batch). While it is unclear in Chapter 3 as whether the random division of trips into sub-steps, or the sequence of vehicle assignment, would affect the simulation outcomes significantly. From the analysis in this chapter, it can be seen that in the case of perfect or nearly perfect information (probe penetration rate $\geq$ 10%), the variabilities in the results are almost undetectable (shown by the data points on the right side of Figure 4.11). Similar sub-stepping process in other studies involves only 4 sub-steps (assigning 40%, 30%, 20% and 10% of the trips in a larger time step), which was adopted in the initial stage of this study. Although the numbers of sub-steps required ultimately depend on the complexity of the network and travel demand, 20 sub-steps are found to be sufficient to stabilise the traffic assignment process and can produce consistent results regardless of the random assignment order. As a result, this sub-stepping process is utilised in the analyses of the coming chapters.

The validity of assuming an hour-long time step

It is a limitation of the mesoscopic traffic simulation model that it currently cannot consider residual demand carried over from vehicles that cannot finish their journeys in a specific time step, mainly due to the computational difficulties. Suppose that there is a larger network where many trips last for around two hours. Then the assumption of hour-long time step without considering the unfinished half of the journeys becomes problematic. One issue associated is that for the unfinished part of the trip, the actual congestion level is lower, meaning that other travellers utilising this part of the network can go even faster. This requires either microsimulations (second by second vehicle movements) or macroscopically, using many iterations to make sure that the final experienced and the simulated journey durations converge. Both solutions are computationally intensive and may not be necessary for certain applications. In SF, an hourly-long time step is a good balance of the computational effort and the temporal resolution. And as indicated by Figure 4.6 and 4.10, the average journey time in the most congested hour (6 PM on a typical Friday) is less than one hour. So an hour-long time step is a valid assumption adopted for this study.

Chapter 5

Pavement degradation modelling

5.1 Overview

Well-maintained pavement infrastructure has great benefits to the transportation system, among which the reduced fuel consumption and carbon emissions are noticeable ones. Driving on smoother road reduces the rolling resistance, thus saving more energy and producing less CO_2. As the aim of this thesis is to study the carbon mitigation strategies of an urban transportation system, the dynamics of pavement condition change over time as well as the space dimensions are included in the overall framework.

The focus of this chapter is to propose a pavement degradation model. The projected pavement condition data will be used in the final chapter to adjust the driving costs and emission factors. In addition, based on the projected pavement conditions, various maintenance strategies will be tested for their impacts on the transportation system carbon emissions. Statistical models of pavement degradation are notoriously hard to build. First of all, unlike traffic, large scale data related to pavement degradation are not abundant. This is a major limitation for many past studies, thus many of the previous works can only focus on small-scale road-level analysis. Secondly, even in cases when large databases dedicated to pavement condition data are available, there are still important factors missing, such as the construction quality, weather conditions and material quality. In addition, the assessment of pavement condition is an error-prone process, depending largely on the subjective judgements of survey personnels. In many ways, the analysis in this chapter is also affected by these constraints. However, a spatial pavement degradation model inspired by methods in geographical studies is proposed to take advantage of the spatial structures in the urban pavement network and to overcome the above constraints to a certain extent (e.g., missing data in parts of the network, measurement errors and etc.). The spatial model is compared with two non-spatial models to assess their performance in pavement degradation modelling.

5.2 Framework

The structure of this chapter is illustrated in Figure 5.1. The real pavement condition degradation dataset is obtained for the city of San Francisco (SF). This dataset is based on more than two decades of pavement condition surveys conducted by qualified personnel. However, as many real world datasets, the SF pavement condition data are not perfect. First of all, it contains some inconsistencies due to, e.g., missing records. A data cleaning step is adopted to correct these inconsistencies as best as possible for the statistical modelling process later. Also, some factors related to the description of the pavement degradation process are not presented in the dataset, such as the pavement structures or the traffic loading conditions. It is understandable for these factors to be missing, as they are hard to obtain in practice. To address this problem, data from other sources are utilised to supplement the pavement degradation observations, including the passenger car traffic (from Chapter 3 of this thesis) and bus traffic open data. This added information will be used to refine the statistical model of the pavement degradation process.

A variety of model structures are then tested, including the traditional categorical-based models, the extreme case of individual road based models as well as models explicitly taking advantage of the spatial adjacency relationship in the pavement condition dataset. The differences between these model structures can be understood to lie in the "spatial resolution" being used. For example, in the simplistic case of the non-spatial categorical model, pavement condition data are grouped into a few categories. A degradation curve is then fitted for each of these categories. In this case, there are relatively abundant information for building the degradation relationship for each pavement category. However, this model structure does not allow capturing the variabilities within the pavement category and it is unlikely that one single relationship would be sufficient to describe the degradation trend, given the complexity in the actual pavement degradation process. To address this issue, a second model with finer categories is also tested. In this model, each pavement segment itself is regarded as one category and a degradation curve is obtained for each street segment. An obvious problem of this fine granularity is the availability of data, as on average there are only five observations of pavement conditions per street segment over the two decades of data collection period. This prompts the application of the third model, which is still based on individual pavement segments. However, information from neighbouring pavement segments are borrowed when the data for one street are insufficient. This is based on the assumption of spatial proximity that streets closer to each other are more likely to share similar degradation behaviours.

The statistical regression is carried out based on Bayesian inference. In the end, the suitable model structure and parameter set are selected according to several evaluation

metrics, including the fitting and testing errors based on the observation data, as well as the potential link to the traffic degradation models.

5.3 Background

Pavement asset managers have always been seeking methods that help with the decisions of when and where to carry out maintenance [79, 86, 94, 246, 253]. In the past, such decisions were largely hindered by the scarcity of data: usually pavement performance models or insights were based on data collected at a small scale and thus not representative enough given the natural variability of the pavement degradation process [121, 169]. The situation has improved recently as, in many places, pavement inspections are carried out more frequently system-wide [100, 152, 201]. However, when it comes to maintenance planning, there are still many difficulties in producing a reliable pavement condition prediction model, particularly with the strong presence of measurement errors inherent to visual surveys and the lack of knowledge on crucial degradation-affecting factors (e.g. construction quality, history of minor maintenance activities), as encountered in this study.

As a result, to address the issue of "imperfect data", additional structures in the data should be considered as useful information, which will hopefully bring about more insights. There have been several studies incorporating the underlying hierarchies of the pavement degradation process. For example, a model is proposed by Anyala et al. [7] to assess the impact of climate change on pavement rutting. In this hierarchical Bayesian model, level 1 parameters govern the degradation process of each surface group while at the same time being constrained by level 2 parameters (network level). This hierarchical structure is used to reduce the parameter estimation uncertainties. In another study [4], roughness (measured by the International Roughness Index, IRI) was modelled in a linear hierarchical manner, reflecting the structured variations of pavement IRI by each section, highway and road class. These existing studies mainly rely on the known hierarchical structures of the street network as additional information, while in this chapter, it is shown that the similar hierarchical modelling approach can be applied in a more general manner, taking advantage of the natural spatial structures of the street network.

5.4 Data

SF is again the case study area for building the pavement degradation model. Data used for modelling pavement degradation in SF come from two parts: the pavement condition data from the SF government's open data portal, as well as the traffic-related data (traffic load)

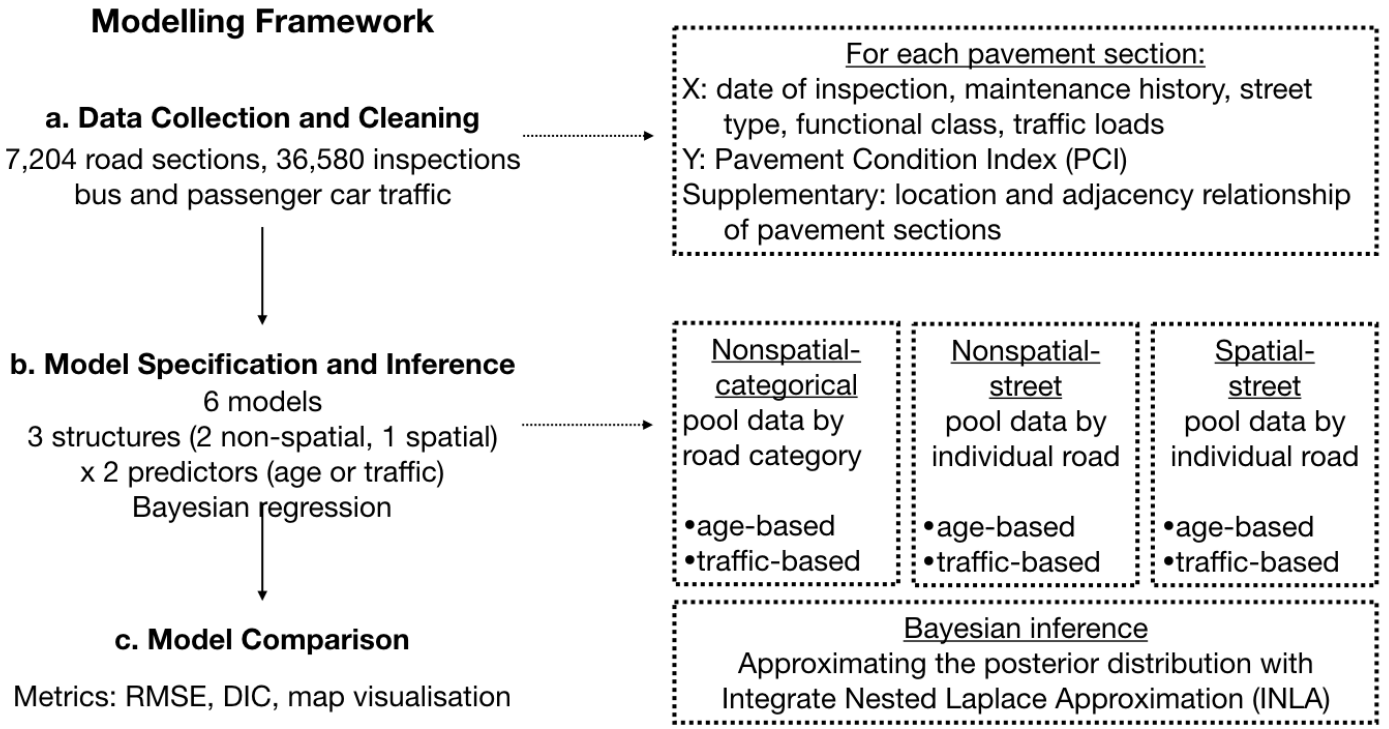


Fig. 5.1 Structure of the chapter.

that have been obtained from the mesoscopic traffic simulations presented in Chapter 3 of this thesis.

5.4.1 Pavement Condition Data

SF Public Works publishes the pavement condition data on DataSF (data.sfgov.org) under the Open Data Commons Public Domain Dedication and License [171]. It provides historical and current information on the Pavement Condition Index (PCI) of more than 12,000 street segments in the city (Figure 5.2). PCI is a numerical scale from 0-100 that is used to represent the general condition of pavement, with 0 being badly deteriorated roads and 100 representing brand new conditions. It was originated in the US in the late 1970s and is still widely used for pavement condition assessment [216, 215]. Other measures of pavement conditions include the IRI, crack rates, rutting depth, etc. However, the PCI condition dataset is the only measure of pavement degradation that has been found for the case study area.

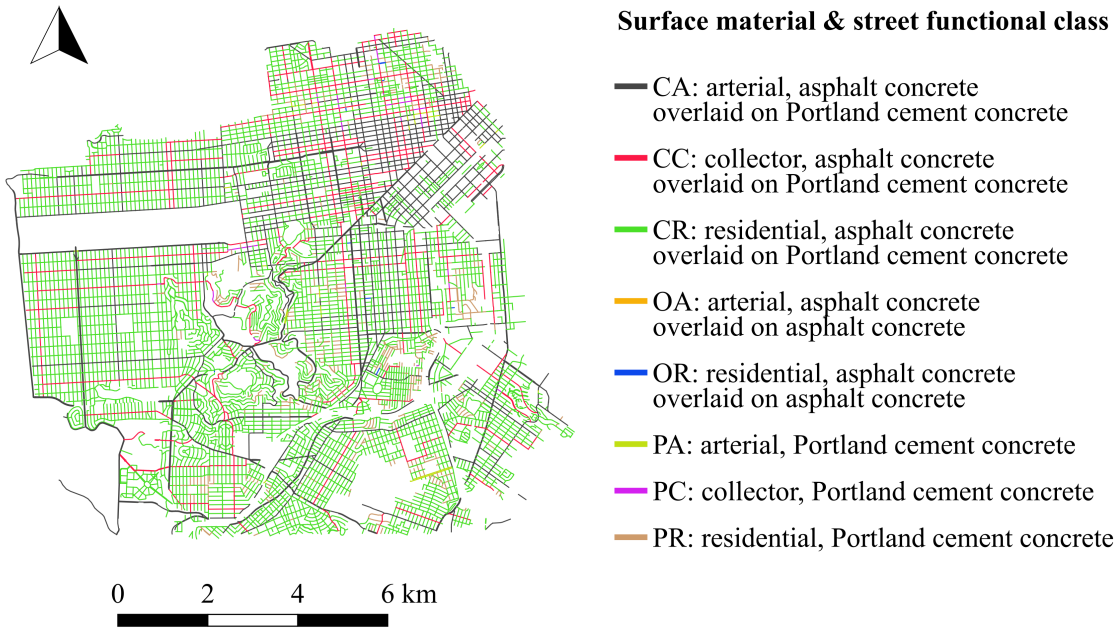


Fig. 5.2 Street network in SF, coloured by surface type and functional class categories.

Table 5.1 offers a glimpse of the dataset by showing the records belonging to street segment "CNN100000" measured at different times over the past 20 years. The pavement condition data are collected by the SF Public Works using visual surveys and they have been used for asset management, decision making as well as publicity purposes. For example, they

are used to demonstrate the pavement condition changes on a yearly basis and to compare with other cities and counties in the Bay Area [235], to assess the outcome of major infrastructure investments [212] and so on. As for pavement performance forecasting, the data are most notably used in the calibration of the pavement performance model in StreetSaver, a Pavement Management Software (PMS) developed by the Bay Area Metropolitan Transportation Commission (MTC) and used by many local agencies in the west coast of the US. The StreetSaver model utilizes a family of deterministic S-shaped curves to predict pavement deterioration as a function of time, with model parameters obtained from weighted least square regression [65]. To further improve the deterministic model in StreetSaver, a stochastic model that projects pavement conditions as a probability distribution is proposed [192]. However, both of these two existing studies categorise the city-scale data by pavement types and do not consider the possible spatial correlations between individual pavement sections.

Table 5.1 Pavement condition records of a street segment

CNN	Street Name	From Street	To Street	Functional Class	Surface Type	PCI Score	PCI Change Date	Maintenance or Survey
100000	01ST ST	Market St	Stevenson St	Arterial	C	100	05/09/2001 12:00:00 AM	Treatment
100000	01ST ST	Market St	Stevenson St	Arterial	C	100	11/19/2002 12:00:00 AM	Survey
100000	01ST ST	Market St	Stevenson St	Arterial	C	100	11/16/2005 04:50:06 PM	Survey
100000	01ST ST	Market St	Stevenson St	Arterial	C	100	07/26/2007 04:21:27 PM	Survey
100000	01ST ST	Market St	Stevenson St	Arterial	C	93	08/10/2009 03:43:09 PM	Survey
100000	01ST ST	Market St	Stevenson St	Arterial	C	73	12/23/2010 01:37:36 PM	Survey
100000	01ST ST	Market St	Stevenson St	Arterial	C	56	01/08/2013 03:18:02 PM	Survey
100000	01ST ST	Market St	Stevenson St	Arterial	C	61	11/25/2014 02:17:39 AM	Survey

Although the earliest record date is in 1947, most record dates are after 1992 (Figure 5.3(a)). Initial explorations of the dataset also show that the PCI records before 1995 may not be fully reliable, since for more than 70% of the street segments, the PCI values from 1992 to 1994 are exactly the same (see an example of such duplicates in Figure 5.3(b)). As a result, only data collected in and after 1995 are used for further analysis.

Ageing is a major factor that leads to the degradation of pavement conditions [180]. In the SF PCI dataset, only kerb-to-kerb maintenance projects are recorded. Based on these maintenance records, pavement "age" since the last maintenance is calculated and used as one choice of the explanatory variables in degradation models. As only maintenance records can help to determine the "age" of the pavements, PCI records without clear previous maintenance dates are thus removed. A scatter plot of pavement segment PCI versus age based on the whole dataset at this stage is given by Figure 5.4. As there are overlapping data points, all data points are made semi-transparent. Thus, the darker the colour, the more points are located at a position. The downward pavement condition degradation trend can be seen in Figure 5.4. However, it is also recognised that the data is very "scattered". For example, pavement conditions at 10 years after maintenance (age = 10) range from about 50 to 90, which indicates a big variation of possible pavement conditions given the total range of the PCI is from 0 to 100.

Next, some obvious outliers of the PCI records are filtered out. Outliers are defined as survey records with an annual change of PCI larger than 40 (Figure 5.5(a)). Besides, as only kerb-to-kerb maintenance works are documented, it means that small scale road works, such as patching or pothole filling, are not reflected in the dataset. To mitigate the influence of these missing maintenance records on inferring the degradation rates, streets that show significant improvements (change of PCI larger than 20 per year) without maintenance have their conditions shifted back to their previous values (Figure 5.5(b)). This allows the examination of the general degradation trend of a particular road segment between major maintenance events. Survey errors and the absence of some maintenance records are the two major limitations in the data. These are actually two pervasive issues in pavement condition databases and need to be solved, e.g., through automated pavement condition surveys or better documentation of road works, for better degradation analyses and pavement asset management purposes.

Traffic, material, climate and construction quality also play important roles in the pavement degradation process [80, 158]. However, not all of these data are well documented or easily accessible. In this study, only the pavement material types and road functional classes are available from the original dataset. Table 5.2 summarises the number of street segments in each material and functional class category. It can be seen that over 97% of the streets

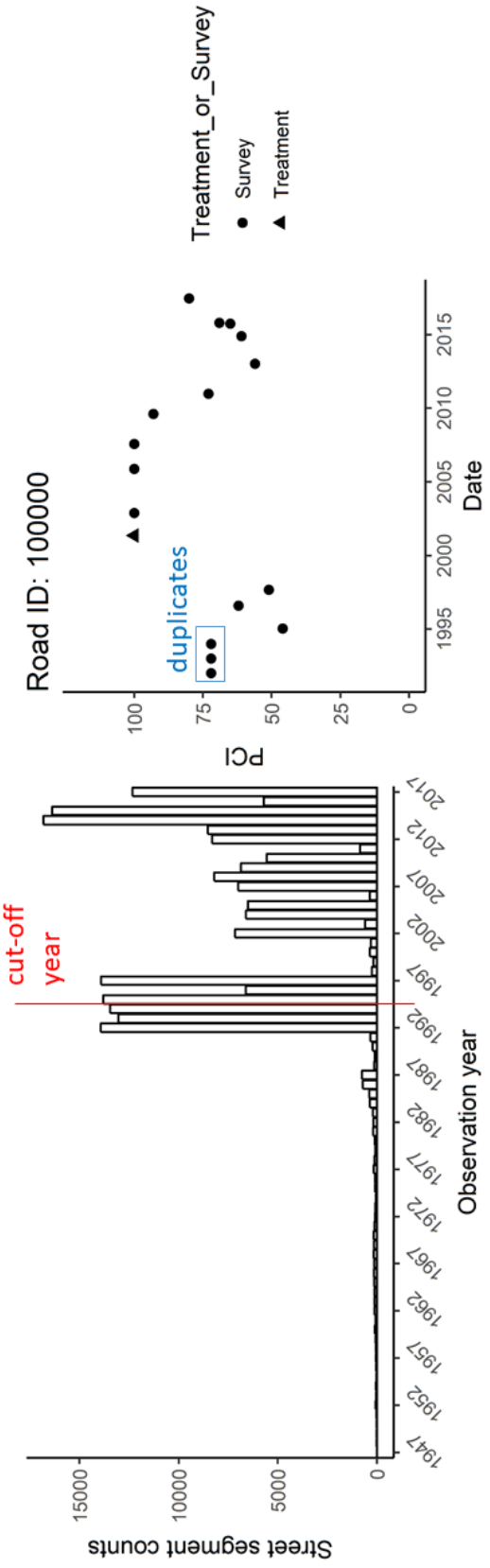


Fig. 5.3 PCI records. (a) A histogram of the observation dates. (b) An example of duplicated PCI values in 1992-1994.

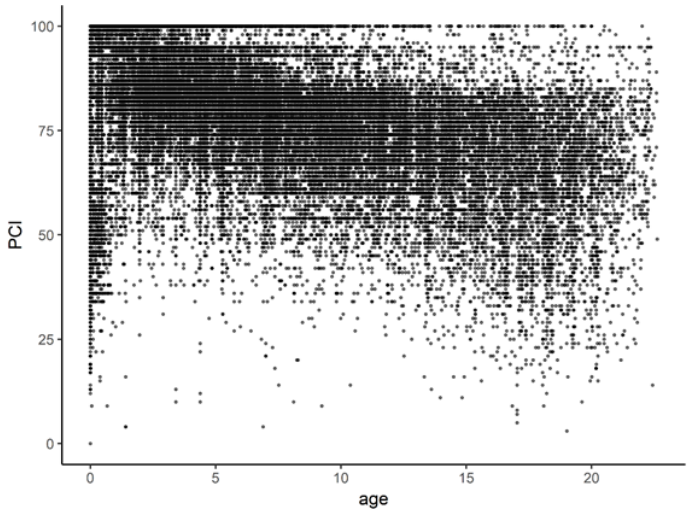


Fig. 5.4 A scatter plot of pavement section's PCI versus age.

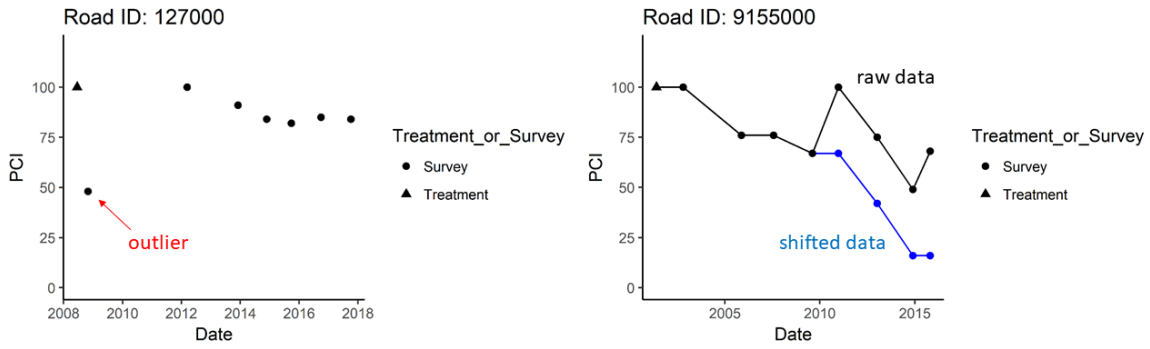


Fig. 5.5 Data cleaning examples. (a) Removing outliers. (b) Handling potential missing maintenance records.

are asphalt concrete overlaid on top of Portland cement concrete. For the functional class classifications, 28% streets are classified as arterial roads, 8% as collectors (less important than arterial streets) and the rest 64% are residential roads. These categorical characters will be taken into account in the subsequent degradation modelling.

After the above basic data cleaning and processing, 8,298 street segments and 55,111 PCI records remain in the dataset. This equals to about 6.64 observations per street segment during the study period from 1995 to 2017. However, this is only the first step of data cleaning to filter out the obvious problematic data points. A further step of processing needs to be conducted while merging the pavement degradation data with the traffic data.

Table 5.2 Pavement categories and numbers of street segments in each category

Street segment counts (in brackets: abbreviations of category types)	Arterial (A)	Collector (C)	Residential
Portland cement concrete (P)	99 (PA)	82 (PC)	1192 (PR)
Asphalt concrete overlaid on asphalt concrete (O)	8 (OA)	NA	40 (OR)
Asphalt concrete overlaid on Portland cement concrete (C)	13817 (CA)	3780 (CC)	30524 (CR)

5.4.2 Traffic Data

As shown in the literature review in Chapter 2, pavement age and traffic load are two most widely used explanatory variables to model the pavement degradation process. As no information about the typical or cumulative traffic load is available from the SF pavement condition dataset, traffic volume results from the mesoscopic traffic simulation in Chapter 3 (or the special case in Chapter 4 with 100% probe penetration rate and no probe data variability) are used to augment the pavement degradation modelling dataset. Specifically, the traffic volumes for a typical week in 2016 are calculated and this weekly traffic is scaled to cover the whole period of the pavement degradation data from 1995 to 2017 by using the bay area traffic growth rate. In addition, bus volumes are also included, while trucks are not, as the road network within the jurisdiction of SF Public Works is not part of the known truck routes (heavy trucks use designated routes that are within the jurisdiction of another agency, Caltrans). The passenger car traffic volumes from the mesoscopic traffic simulation and the bus volumes are then combined to calculate the cumulative equivalent single axial load (ESAL, a measure of pavement damage potential) since the last maintenance for each pavement condition inspection data point.

The passenger car traffic volume

The traffic simulation in Chapter 3 is based on the travel demand of passenger vehicles. Pavement damages caused by these small vehicles are very often ignored, as "even a fully loaded large passenger van will only generate about one-thousands of the damage a fully loaded tractor-semi trailer can generate" [184]. However, as the residential traffic constitutes the majority of all types of traffic in the local road network, it is still good to check the impacts of passenger vehicles in this study. Figure 5.6a shows the weekly traffic (including

weekdays and weekends) from the traffic simulation. The original travel demand for the traffic simulation is based on data collected in 2016, while the pavement condition data started in 1995. As no information has been found regarding the total traffic growth in the city of SF in the past two decades specifically, the growth trend of traffic entering and leaving the whole Bay Area at regional gateways is used as a proxy [236]. It can be seen in Figure 5.6b that this gateway traffic grew steadily from the mid 1990s till the mid 2000s and start to decline around 2007, possibly influenced by the financial crisis. The decreasing trend started to reverse around 2012. The series stops in 2016, so the traffic of the year of 2017 is projected by assuming the same growth rate from 2015 to 2016. An estimation of volumes of the passenger cars on the SF roads can thus be obtained based on these scaling ratios as presented in Figure 5.6b.

One issue encountered at this step is the transferring of information from the traffic simulation model to the pavement degradation model as the inputs. The traffic simulations are conducted on the highly detailed OpenStreetMap (OSM) network, while the pavement condition data are provided on a simplified network Figure 5.8 (a) & (b). There is no common identification of streets across these two networks. As a result, a spatial conflation process is carried out so as to obtain a mapping relationship of the two networks:

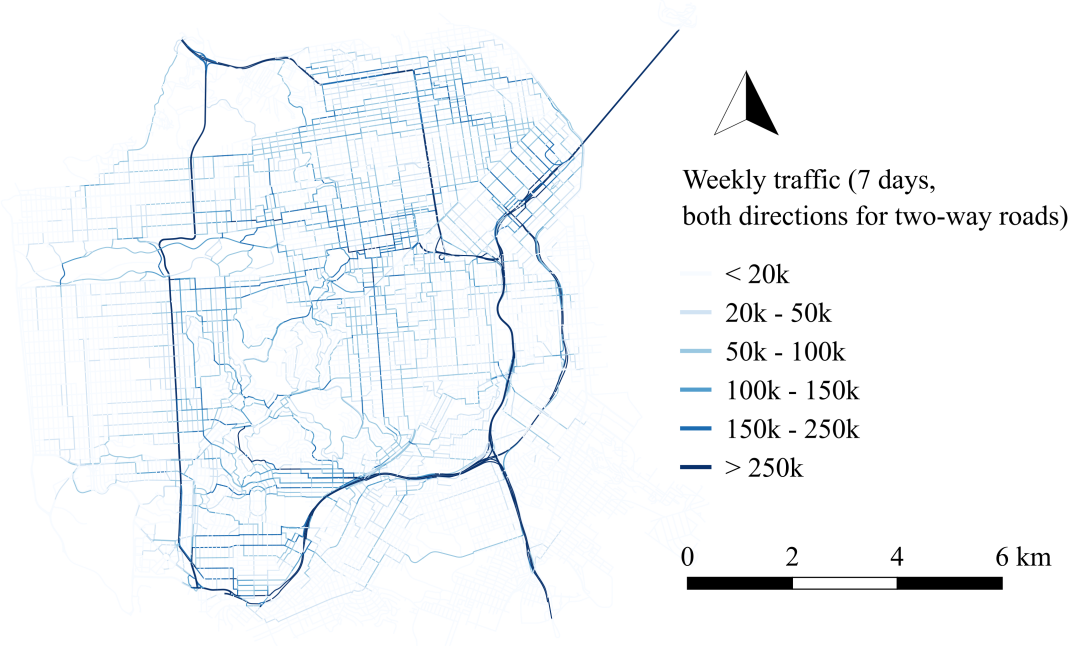
1. For a pavement link, draw a buffer of 5×10^{-5} degrees. This corresponds to the physical distance at about several metres and is found to provide a good buffer size for this specific conflation task through trial-and-error. This is shown schematically as the blue region around ”pavement link 1” in Figure 5.7 (a).

2. Find all traffic links that intersect with this buffer region. Tagging these traffic links as ”potential matches”. In Figure 5.7 (a), only the red ”traffic link” proves feasible. This step uses R-tree spatial indexing inspired by Boeing [26].

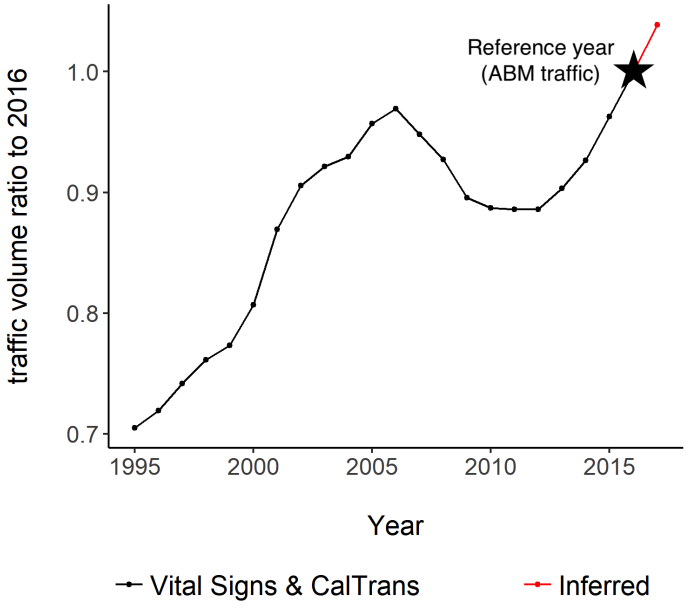
3. Calculate the intersection length between all ”potential matches” and the buffer zone.

4. Find the ”exact match”, which should satisfy the following three criteria simultaneously:

    - The intersected length should be no shorter than 90% of the maximum intersected length of all ”potential matches”. This rule is set to retain the most likely matches among all the ”potential matches”;
    - The intersected length should be longer than 20% of the total length of the same traffic link. This is to rule out cases when the traffic link touches the pavement link, but not parallel to it, as shown in Figure 5.7 (b).



(a) Weekly passenger car traffic volume from the mesoscopic traffic simulation.



(b) Bay Area gateway traffic growth at regional gateways [236].

Fig. 5.6 Passenger car volume.

- Among all the remaining candidate traffic links, the final choice is the link whose intersection length is the closest to its total length.

5. Repeat from Step 1 till there is a matching traffic link for all the pavement condition link.

6. After the above automatic procedure, a visual inspection is conducted around areas with complex geometry to correct any obvious mismatches.

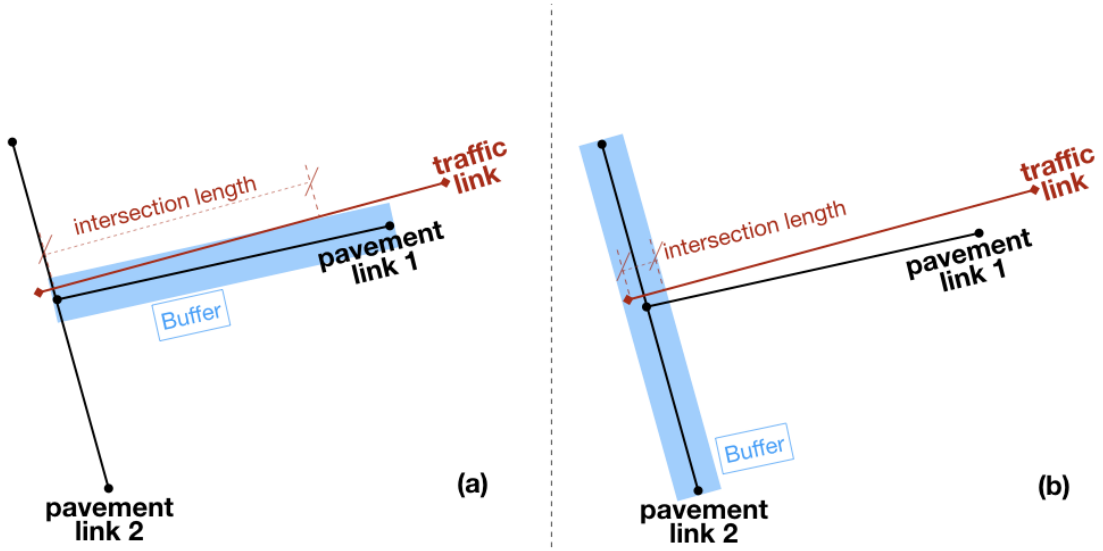


Fig. 5.7 Illustration of the graph conflation process. (a) The traffic link has large intersection length with the buffer of pavement link 1. (b) The traffic link has small intersection length with the buffer of pavement link 2.

As there is no fixed formula or a general solution for graph conflation, the above steps (including the buffer size) are devised based on a trial-and-error process. These procedures prove to work well for this case. The performance of the matching algorithm is manifested by Figure 5.8 (c) & (d). The black links are those storing the pavement condition information (pavement links) and the black numbers are the unique index for the pavement links. The red links are those storing the traffic amount data (traffic links), while the red numbers are the corresponding index of their matching pavement links. It can be seen that nearly all the traffic links and pavement links are matched correctly (the black number being the same as the red number). This establishes the correspondence between the traffic simulation network and the pavement network and allows the transfer of data from one to another.

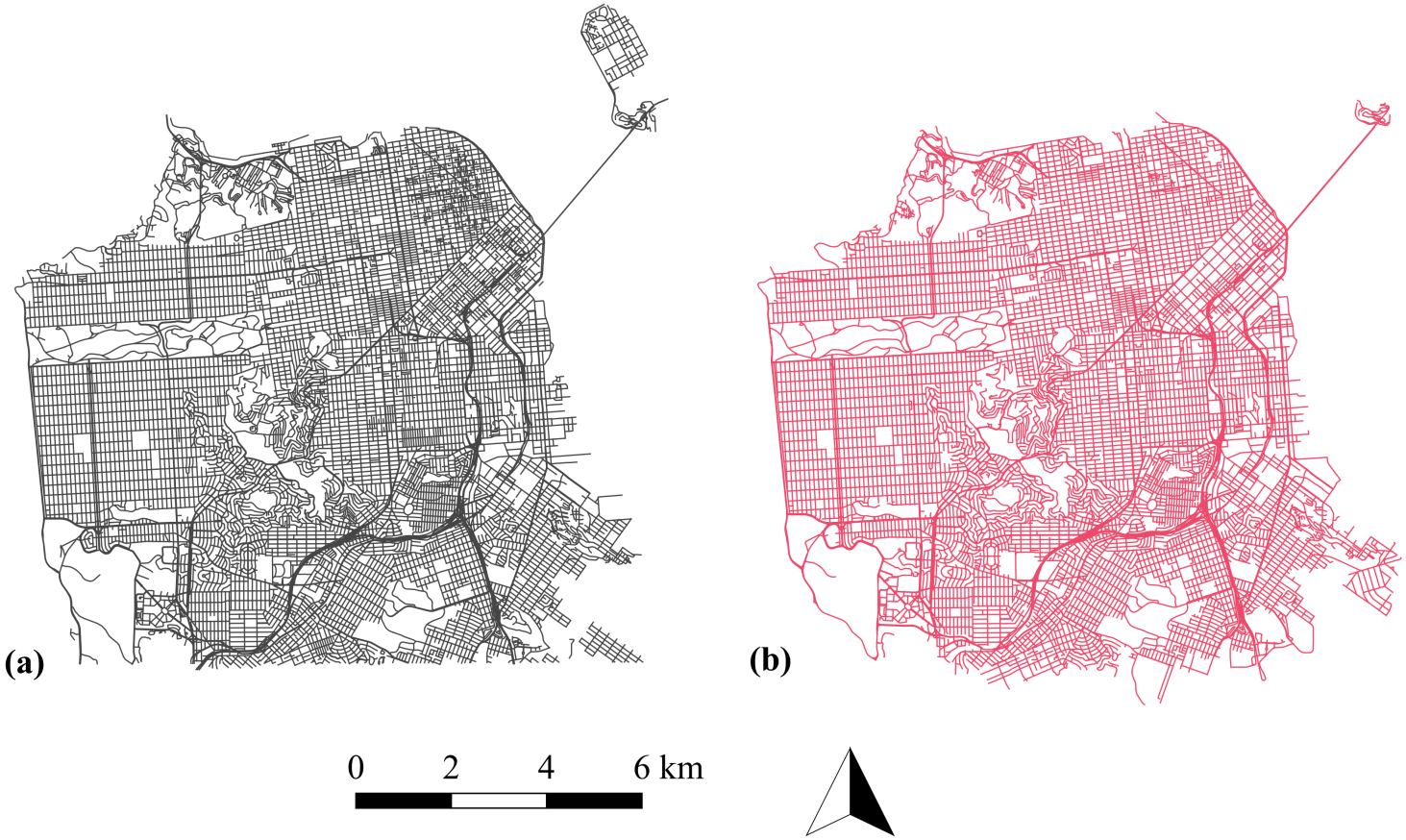


Fig. 5.8 Merging information from the traffic map and the pavement condition map (to be continued on the next page). (a) Pavement condition map (from the SFDPW); (b) Traffic simulation map (from the OSM). These two maps are not identical.

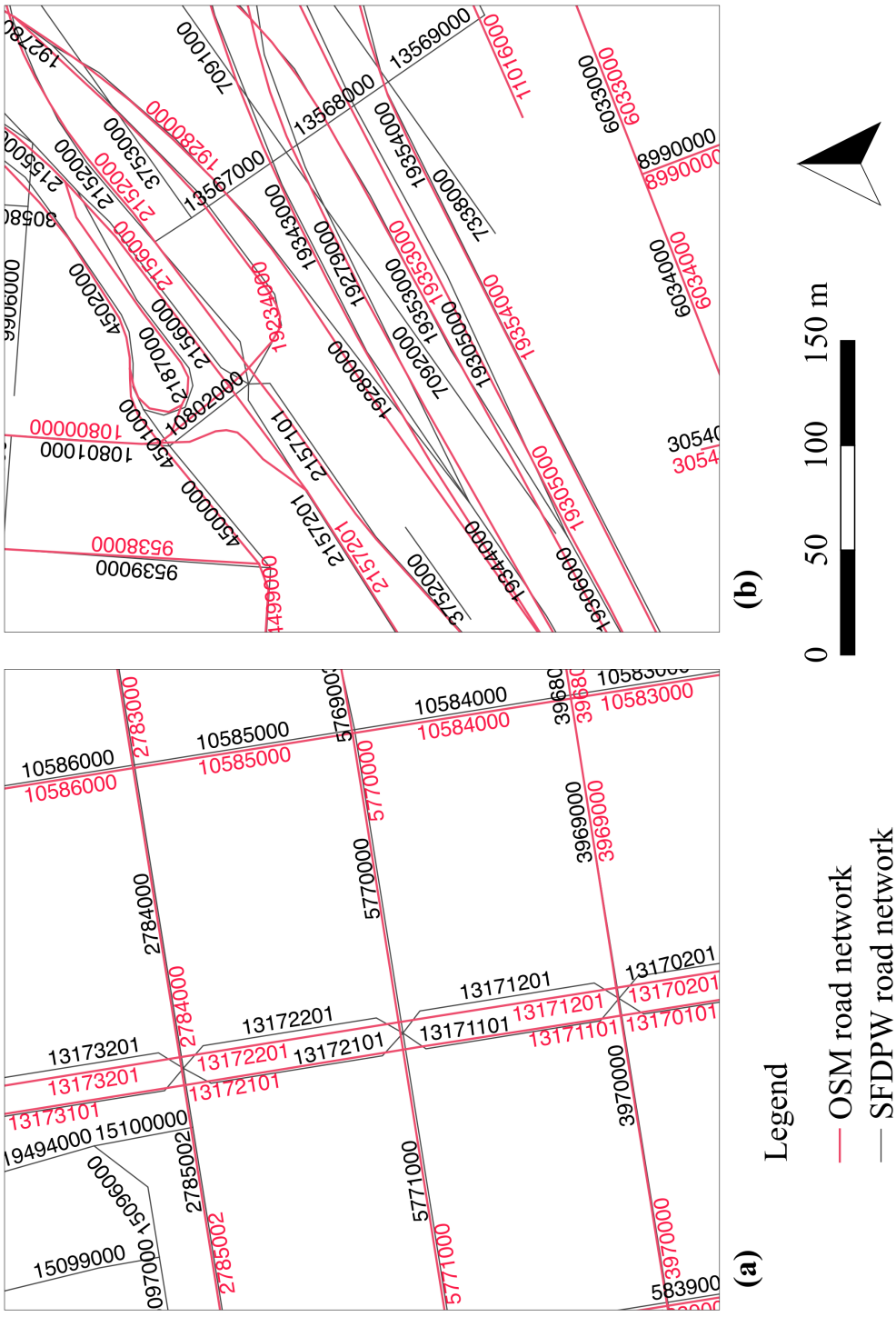


Fig. 5.8 Merging information from the traffic map and the pavement condition map (cont.). (c) & (d) Zoom-in views showing the discrepancies between the two maps and conflation performance.

The bus and truck volume

Designated truck routes in SF are obtained from the Caltrans website and are compared with the SF Public Works network (Figure 5.9(a)). It turns out that none of the road segments in the SF Public Works dataset belongs to the truck routes. This is sensible as the SF Public Works is a local agency and is not responsible for the maintenance of highways. As a result, it is assumed that no significant truck volume exists on the pavement network in this study. The highways are not considered in the pavement degradation modelling as their structures are different from the local roads and are usually kept in good conditions. But the highways are an important component for the mesoscopic traffic simulation part of this thesis.

Even though the trucks are not considered, the buses (which have a special name in SF, the Muni) do operate on the local network. As a result, the Muni routes and schedules are downloaded from the SF Municipal Transportation Agency (SFMTA) website. The data is provided in the General Transit Feed Specification (GTFS) format, which was devised by Google in 2005 to facilitate the sharing of public transit schedules. The quantity of interest is the weekly bus volume, which can be obtained by grouping together the bus trips that share the same path geometry (*shape_id* in GTFS). Then, this path geometry is mapped to its nearest pavement segments in a similar process shown in the spatial conflation previously in Figures 5.8 and 5.7. Figure 5.10 shows the processing procedure of the GTFS data and Figure 5.9(b) is the weekly bus volumes mapped to the SF Public Works pavement network.

The bus routes are mapped to about one third of the pavement segments. The segments of Market Street between the 1st and 2nd Street in downtown SF have the highest bus volume. Each week 15,924 buses pass by, about 1.5 per minute. This is not surprising as there are 19 bus lines operate on this part of the street. On average, the weekly bus volume is found to be 1,526 on the subset of pavement segments serving buses.

Converting the traffic volume to Equivalent Single Axial Load (ESAL)

The weekly passenger car volumes and bus volumes on the SF road network by themselves are not the best variables to be used for predicting pavement degradations, as it is the vehicle weight, fleet composition, pavement type and structure that determine the damage caused by the traffic. So the traffic volumes obtained above first need to be converted into a damage potential measure.

A widely used approach is to convert the damage caused by a specific vehicle to the damage caused by a standard load, e.g., the 18 kips (80 kN) ESAL [184]. According to the Caltrans Highway Design Manual, the ESAL of a 2-axle bus is around 0.378 [38]. An 0.001 ESAL is considered for passenger car. Moreover, the vehicle distribution to dedicated bus

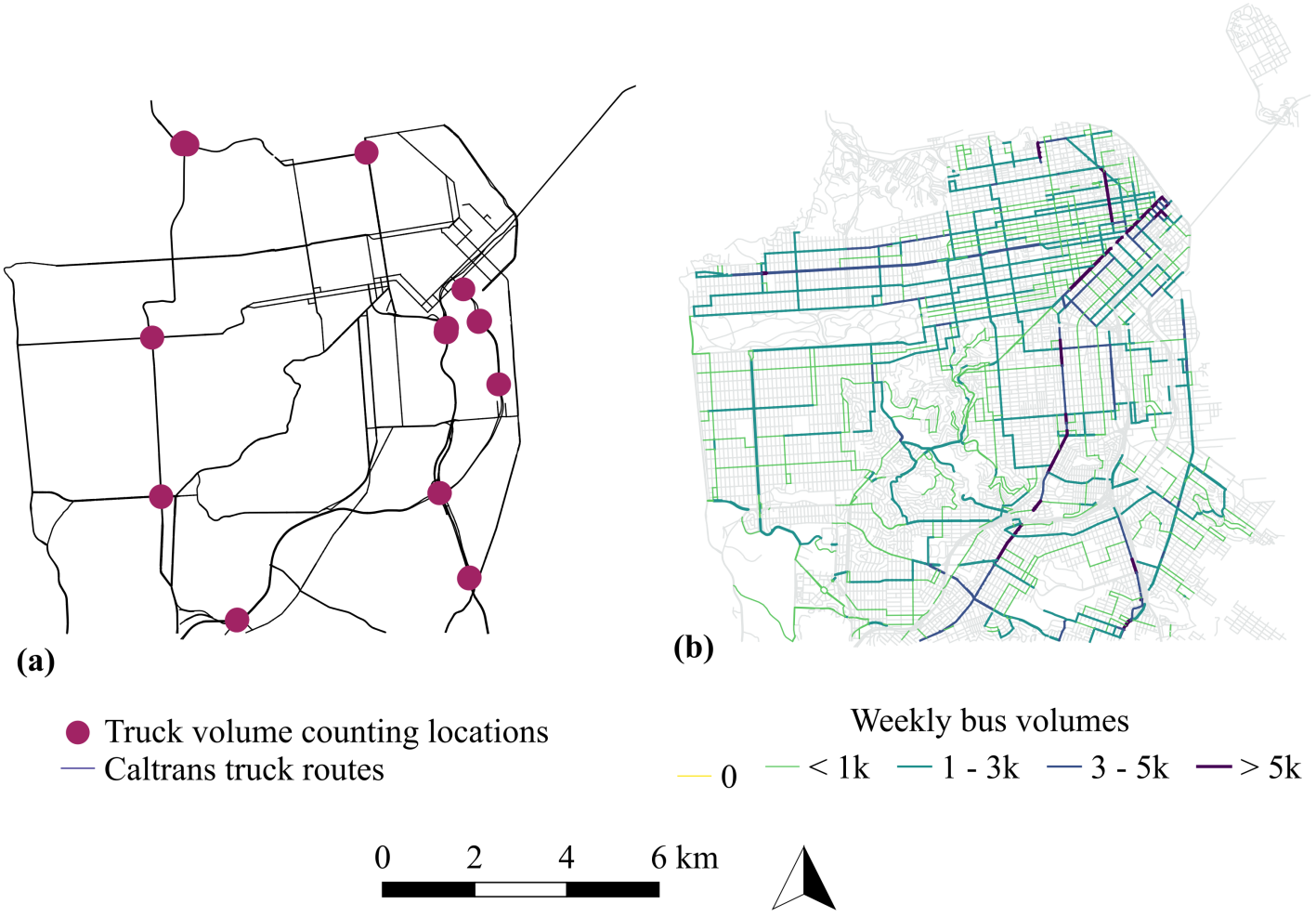


Fig. 5.9 Truck and bus routes in SF. (a) Truck routes from Caltrans (not overlapping with the SF Public Works roads). (b)Weekly bus volume from the SFMTA.

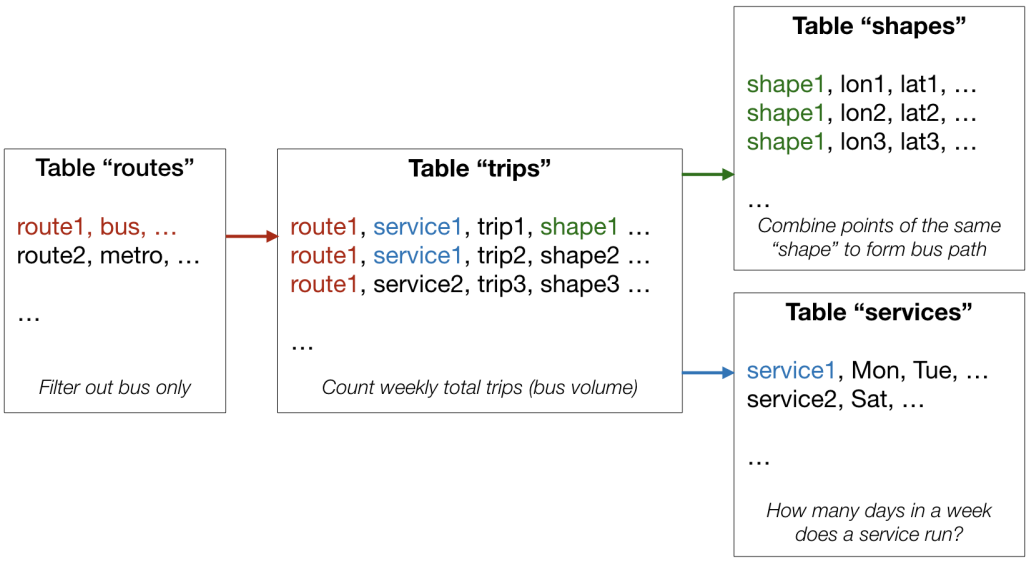


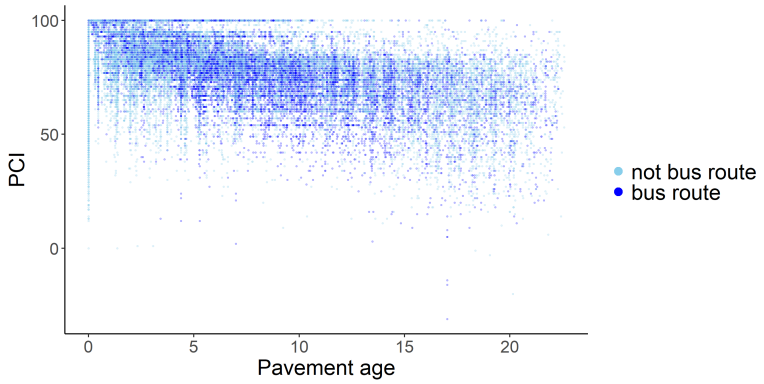
Fig. 5.10 Processing of SFMTA GTFS data to get weekly bus volumes.

lanes or High Occupancy Vehicle (HOV) lanes are not considered, which is reasonable for this study as there is no information as for in which lane(s) the PCI data are measured, either. The total weekly ESAL for each link is thus calculated as:

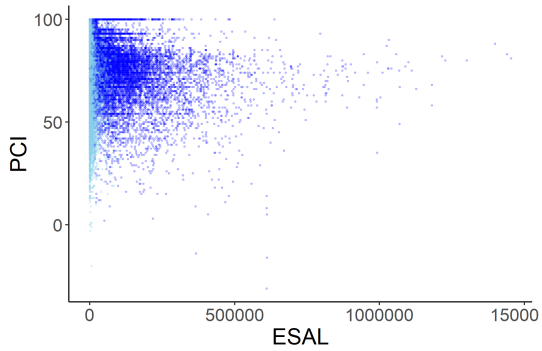
$$weekly\_ESAL = \frac{1}{Number\ of\ Lanes} \times (0.378 \times weekly\_bus\_volume + 0.001 \times weekly\_car\_volume) \tag{5.1}$$

Based on the same traffic growth trend in Figure 5.6b, the total cumulative ESAL since the last maintenance is then calculated for each pavement inspection record. Figure 5.11 plots the change of PCI over time with the pavement age and cumulative ESAL. However, as the distribution of ESAL concentrates on the smaller side, the PCI trend with log scale of the cumulative ESAL is plotted as well. In Figure 5.11 (c), the cumulative ESAL values are shifted to the right by 1, so that the log values are non-negative, as the horizontal axes in the other two plots. The scatter plot is coloured according to whether the pavement segments carry any bus load. It can be seen that the pavement segments with bus traffic have distinct degradation trends from those without, especially in the bottom two plots where the ESAL or log ESAL is used as the x axis. As a result, the bus/no bus distinction is considered alongside the pavement street type and functional class categories in Table 5.2 in the analysis. Plots in Figure 5.11 are especially useful for determining the explanatory variables for building the

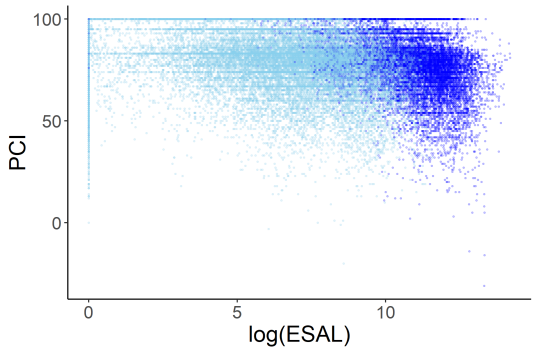
pavement degradation model. Visually, age seems to be the best choice as the trend of PCI degradation with pavement age is the clearest (Figure 5.11(a)). This is followed by log ESAL (Figure 5.11(b)). The PCI against ESAL without log transformation is too scattered (Figure Figure 5.11(c)). In the next section, pavement age and cumulative ESAL will be used as two choices of the explanatory variables for predicting pavement degradation.



(a) PCI change with age



(b) PCI change with cumulative ESAL



(c) PCI change with log cumulative ESAL

Fig. 5.11 PCI degradation trends with age, ESAL and log ESAL, ESAL shifts right by one to set the minimum log ESAL at 0.

In all three plots in Figure 5.11, it can be seen that there is a string of data points at the leftmost vertical line, corresponding to age = 0 or ESAL = 0. It is unlikely that pavements just after maintenance would fail, so it is assumed that these clusters of points are due to inspection errors or data entry errors. Similarly, there are also points at the horizontal line corresponding to PCI = 100. These points also stand out from the rest of the observations. As a result, datasets in these special clusters (age = 0, log ESAL = 0 or PCI = 100) are not used in the modelling process. After this final filtering, the dataset consists of 36,580 PCI observations on 7,204 road segments, or 5 observations per segment on average. Figure 5.12 shows the spatial distribution and the histogram of the number of available data points per

street. The whole dataset is randomly partitioned into a training set and a testing set. The training set consists of 29,810 records, or roughly 80% of the whole set. The testing set contains the rest of the 6,670 records. In the further analysis, the training set will be used to obtain model coefficients and to evaluate how well the models do in fitting a specific large dataset; while the testing set will be used to check the generality, i.e., how well the models perform when tested on data unseen. Although increasing numbers of the study now partitions the whole data into a training, testing and validation set, the validation set is not adopted in this study due to the relatively small numbers of models to be tested and selected from.

5.5 Methodology

As seen in Figure 5.11, the pavement degradation data are rather scattered. As a result, age or log ESAL itself is not sufficient to describe the variations in the dataset. One approach to account for the variability is to divide the datasets according to some categories based on their surface types and functional classes. Also, from Figure 5.11, further divisions based on whether the pavement carries bus traffic or not is potentially helpful as well. However, as there are only 8 surface type functional class combinations (Table 5.2) and 2 bus traffic levels, the categorisation used in this simple model may not be sufficient to represent the diverse pavement characteristics in reality (e.g., structural type, micro climate, etc.). It is also possible to test the other extreme by considering each street itself as a category and grouping the observed data by street IDs. This allows the individual characters of each street (e.g., climate, geology, construction quality, traffic load, etc.) to be fully captured and represented. However, the numbers of parameters involved in this model will also be large and may lead to overfitting. In this chapter, a third alternative is tested that involves a medium cluster size, smaller than a street type/functional class category, but larger than an individual street. Spatial modelling offers this capability by incorporating dependencies between neighbouring spatial units and is chosen as the third modelling strategy in this chapter.

5.5.1 The spatial model

The spatial model allows smoothly varying coefficients across the entire study area. A simple way to incorporate spatial relationships in a model is to include the longitude and latitude coordinates as model predictors and fit in a trend surface. However, such models can only capture big, global trends if not using high order terms [137]. To represent localised interactions, model structures such as the Geographically Weighted Regression (GWR), Simultaneous

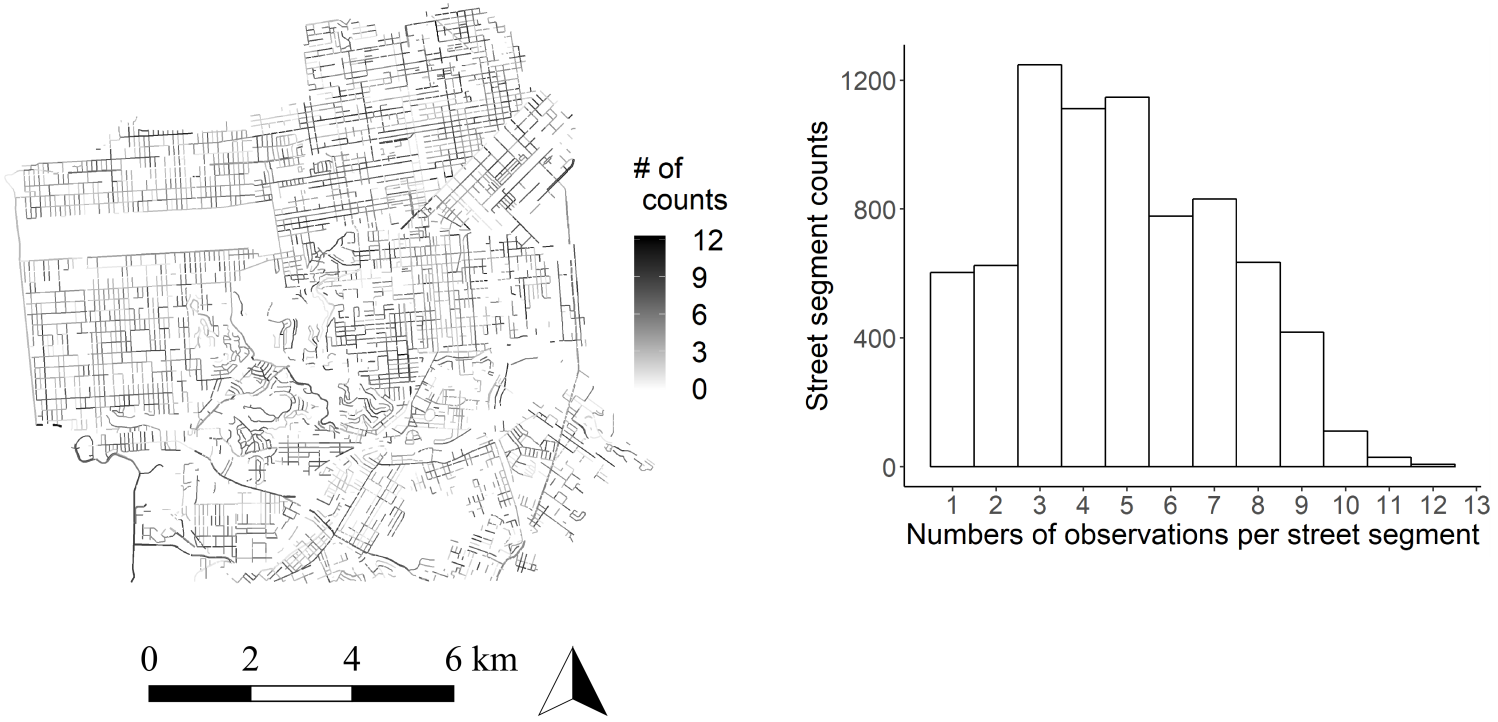


Fig. 5.12 PCI observations per street segment. (a) Spatial distribution. (b) Histogram of the distribution.

Autoregression (SAR) and Conditional Autoregression (CAR) are more suitable [33]. These model structures consider spatial dependencies between neighbours by imposing constraints on the values (or residuals) of neighbouring sites [68]. Readers can refer to Jahanbakhsh et al. [116] for a SAR model that predict pavement conditions with spatial and temporal lags. This study is designed to model street-specific degradation rates as the average of neighbouring rates, so an intrinsic conditional autoregressive (iCAR) 'Besag' model becomes a natural choice [18, 23, 135].

Pavement degradation with spatial effects can be formulated as a multi-level model shown in Equation 5.2 [22, 199]: the first step is to model the distribution of the observations:

$$y_{ik} \sim N(\eta_{ik}, \frac{1}{\tau_g}) \tag{5.2}$$

where y_{ik} is the k-th observed PCI at street i, given that most of the street segments have more than one record in the past 20 years (Figure 5.12). y_{ik} is assumed to follow a normal distribution with mean η_{ik} and precision τ_g (inverse of variance). The distribution equation states that the observed PCI centres around an unobservable mean η_{ik}, plus some random deviations determined by the precision parameter τ_g. In the second step, η_{ik} is modelled as a linear combination of the explanatory variables:

$$\eta_{ik} = \alpha + \xi_i + (\beta + v_i + u_i)x_{ik} \tag{5.3}$$

$$\xi_i \sim N(0, \frac{1}{\tau_\xi}) \tag{5.4}$$

$$u_i \sim N(0, \frac{1}{\tau_u}) \tag{5.5}$$

$$v_i | v_{j \neq i}, \tau_v \sim N(\frac{1}{n_i} \sum_{j \sim i} v_j, \frac{1}{n_i \tau_v}) \tag{5.6}$$

x_{ik} is the explanatory variable (age or cumulative ESAL) corresponding to observation y_{ik}. α and β are the global average intercept and age/traffic load effect shared by all streets. ξ_i is the street specific variation in intercept, which itself is a random variable following a zero-mean normal distribution (Equation (5.4)). $\beta + v_i + u_i$ is the total street specific age or traffic load effect for street i, where v_i is the spatially structured individual deviation from the mean and u_i is the unstructured part. A Besag specification is adopted for modelling v_i [18, 23]: as shown in Equation (5.6), v_i is a Gaussian random variable whose mean equals to the average of neighbouring sites' values v_j ($j \sim i$ means i, j are neighbours) and whose precision τ_v is to be estimated from the data. n_i is the numbers of neighbours that street i has. Since v_i is only related to its neighbours, v_i and v_l are conditionally independent if i and l are not neighbours

(the Markov property). So $\boldsymbol{v} = \{v_1, v_2, ...\}$ is said to be a Gaussian Markov Random Field (GMRF).

In the third step of the multi-level model, prior distributions are assigned to model parameters. Details on prior distributions will be given in the modelling section. A graphical model for the three-level spatial model is given in Figure 5.13(d).

5.5.2 Bayesian regression using R-INLA

For multi-level models such as the one presented above, various techniques are available for parameter inference, including the maximum likelihood estimation (MLE), the Markov Chain Monte Carlo (MCMC) method and so on [59, 91]. Rue et al. [199] has demonstrated that a direct approximation based Bayesian approach, called the Integrated Nested Laplace Approximation (INLA), to be fast and sufficiently accurate for parameter inferences of spatial hierarchical models. In the preliminary stage of this study, the MCMC approach also showed promising and comparable results. Since the scope of this chapter is to compare spatial and non-spatial model structures rather than the various inference methods, only INLA, the fastest and most flexible approach according the author's experience, is adopted.

Continuing with the notation definitions in Equation (5.2)-(5.6), the task of regression is to estimate model parameters $\boldsymbol{\theta} = \{\alpha, \beta, \boldsymbol{\xi}, \boldsymbol{v}, \boldsymbol{u}\}$ and hyperparameters $\boldsymbol{\psi} = \{\tau_g, \tau_\xi, \tau_u, \tau_v\}$ from the data. Based on the Bayes' theorem and conditional probability, the joint posterior distribution of the unknowns given the data $\pi(\boldsymbol{\theta}, \boldsymbol{\psi}|\boldsymbol{y})$ is given by:

$$\pi(\boldsymbol{\theta}, \boldsymbol{\psi}|\boldsymbol{y}) \propto \pi(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\psi})\pi(\boldsymbol{\theta}|\boldsymbol{\psi})\pi(\boldsymbol{\psi}) \tag{5.7}$$

From Equation (5.7), marginal posteriors of a parameter, $p(\theta_w|\boldsymbol{y})$, and a hyperparameter, $p(\psi_h|\boldsymbol{y})$, can be obtained through integration:

$$\pi(\theta_w|\boldsymbol{y}) = \int \pi(\theta_w|\boldsymbol{\psi}, \boldsymbol{y})\pi(\boldsymbol{\psi}|\boldsymbol{y})d\boldsymbol{\psi} \tag{5.8}$$

$$\pi(\psi_h|\boldsymbol{y}) = \int \pi(\boldsymbol{\psi}|\boldsymbol{y})d\boldsymbol{\psi}_{-h} \tag{5.9}$$

INLA does the above integrations through approximating the integrands with known distributions. Based on Tierney and Kadane [228], Rue and Martino [198] proposed the

following approximation:

$$\pi(\boldsymbol{\psi}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{\theta},\boldsymbol{\psi}|\boldsymbol{y})}{\pi(\boldsymbol{\theta}|\boldsymbol{\psi},\boldsymbol{y})} \propto \left.\frac{\pi(\boldsymbol{\theta},\boldsymbol{\psi},\boldsymbol{y})}{\widetilde{\pi}_G(\boldsymbol{\theta}|\boldsymbol{\psi},\boldsymbol{y})}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*(\boldsymbol{\psi})} \tag{5.10}$$

where $\widetilde{\pi}_G(\boldsymbol{\theta}|\boldsymbol{\psi},\boldsymbol{y})$ is the Gaussian approximation of $\pi(\boldsymbol{\theta}|\boldsymbol{\psi},\boldsymbol{y})$ near its mode $\boldsymbol{\theta}^*(\boldsymbol{\psi})$. The formula is equivalent to the Laplace approximation of marginal posterior density in Tierney and Kadane [228]. Similarly, the other integrand in Equation (5.8) is approximated by [199]:

$$\begin{aligned}\pi(\theta_w|\boldsymbol{\psi},\boldsymbol{y}) &= \frac{\pi(\theta_w,\boldsymbol{\theta}_{-w}|\boldsymbol{\psi},\boldsymbol{y})}{\pi(\boldsymbol{\theta}_{-w}|\theta_w,\boldsymbol{\psi},\boldsymbol{y})} \\ &\propto \left.\frac{\pi(\boldsymbol{\theta},\boldsymbol{\psi},\boldsymbol{y})}{\widetilde{\pi}_G(\boldsymbol{\theta}_{-w}|\theta_w,\boldsymbol{\psi},\boldsymbol{y})}\right|_{\boldsymbol{\theta}_{-w}=\boldsymbol{\theta}^*_{-w}(\boldsymbol{\psi})}\end{aligned} \tag{5.11}$$

Substituting the integrands in Equations (5.8) and (5.9) with Equations (5.10) and (5.11), the marginal posterior distributions become integrations at a much lower dimension, which can then be solved numerically. In this study, the R package 'INLA' (www.r-inla.org) is used for the Bayesian INLA regression. Apart from the methodological references by Rue and Martino [198] and Rue et al. [199], information about INLA applications can also be found in Blangiardo et al. [23] and Schrödle and Held [209].

5.6 Models

Three functional forms combined with two explanatory variables (age and log ESAL) are designed to represent an array of modelling strategies:

1. Non-spatial categorical models:

    - $categorical-age$: a non-spatial model with data divided into coarse categories based on pavement surface type, street functional class (Table 5.2) and whether it is part of the bus route. Age is used as the primary explanatory variable for the regression;
    - $categorical-logESAL$: similar to $categorical-age$ except that the explanatory variable for the regression is the log cumulative ESAL;

2. Non-spatial models for individual roads:

    - $street-age$: a non-spatial model with data divided into fine categories based on street ID, with age as the regressor;

- $street - logESAL$: similar to $street - age$, with log cumulative ESAL as the explanatory variable;

3. Spatial models for individual roads:

   - $SP - age$: a spatial model with categories based on the spatial structure of the pavement network, with age as the regressor;
   - $SP - logESAL$: similar to $SP - age$, with log cumulative ESAL as the explanatory variable.

The mathematical expressions for these models are given in Table 5.3. Directed acyclic graphic (DAG) models showing variable relationships are provided in Figure 5.13. For the purpose of clarity, the DAG models in Figure 5.13 are based on a simplified network which consists of only four road segments in two pavement categories (Figure 5.13(a)), as opposed to the 7,204 segments and 14 pavement categories (some combinations of surface type, functional class and bus routes are empty) in the real dataset. Nonetheless, the simplified network and DAG models in Figure 5.13 are sufficient to illustrate the structures and differences of the three models.

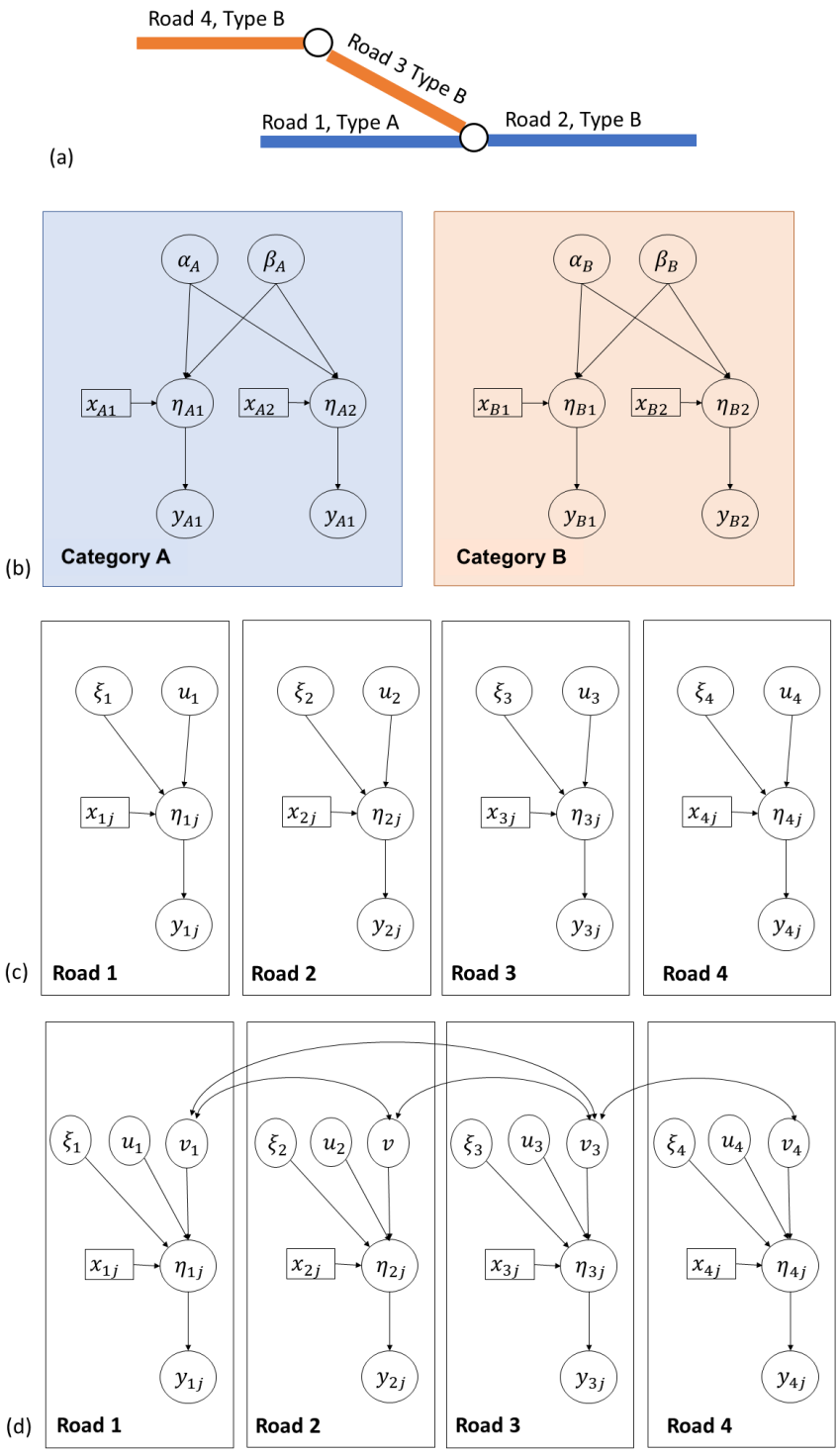


Fig. 5.13 Graphical models of road degradation. Symbols are consistent with definitions in Equation 5.2-5.6 and Table 5.3. For clarity, random noises are not shown in the graphical models. Global-level variables (α, β in the non-spatial individual road based models and the spatial models, as well as the hyperparameters) are not shown, either. (a) An example network made of 4 road segments and 2 road type categories; (b) Non-spatial categorical model: coarse categorisation based on road type category; (c) Non-spatial individual street ID based model: fine categorisation based on individual roads; (d) Spatial model: spatial models with correlated parameters between neighbouring road segments.

Table 5.3 Models

No.	Features	Model	Definitions
$categorical-age$ $categorical-logESAL$	Global level; regression by category	$y_{Ik}=\alpha_I+\beta_I x_{Ik}$	I: index for road type, functional class and bus route category y_{Ik}: k-th observation in category I x_{Ik}: corresponding pavement age or log ESAL for y_{Ik} α_I and β_I: regression parameters for category I
$street-age$ $street-logESAL$	Street level; Spatially-unstructured intercepts and age (or traffic) effects	$y_{ik}=\alpha+\xi_i+(\beta+u_i)x_{ik}$ $\xi_i\sim N(0,\tau_\xi^{-1})$ $u_i\sim N(0,\tau_u^{-1})$	i: index for road segment y_{ik}: k-th observation for road i x_{ik}: corresponding pavement age for y_{ik} α and β: global average intercept and age (or traffic) effect ξ_i: street-level intercept deviation from α u_i: street-level age (or traffic) effect deviation from β τ_ξ: precision for the iid variable ξ τ_u: precision for the iid variable u
$SP-age$ $SP-logESAL$	Street level; Spatially-unstructured intercepts and spatially-structured age (or traffic) effects	$y_{ik}=\alpha+\xi_i+(\beta+u_i+v_i)x_{ik}$ $\xi_i\sim N(0,\tau_\xi^{-1})$ $u_i\sim N(0,\tau_u^{-1})$ $v_i\|v_{j,i\neq j},\tau\sim N(\frac{1}{n_i}\sum_{i\sim j}v_j,\frac{1}{n_i\tau_v})$	i, y_{ik}, x_{ik}, α, β, ξ_i, τ_ξ, τ_u: same as in $street-age$ and $street-logESAL$ u_i: the spatially-unstructured street-level age (or traffic) effect deviation from β v_i: the spatially-structured street-level age (or traffic) effect deviation from β $i\sim j$: i and j are neighbours τ_v: the precision related parameter for spatially-structured random variable v

5.6.1 Non-spatial categorical models

The non-spatial categorical model series is a simple structure where pavement condition observations are divided into categories first based on the surface materials, street functional classes and whether a segment belongs to bus routes, as shown in Figure 5.14 and 5.15. Separately in each category, the condition degradation is modelled with a non-spatial linear form, where age or log ESAL is the main explanatory variable. Visually, this is equivalent to fit a linear trend to data in each cell in Figure 5.14 and 5.15. Although more complex model forms can be used, such as the non-linear trend used by the Bay Area MTC [65], a linear form is a good starting point for showing the overall trend and comparing with the non-spatial individual models and the spatial models, which are also linear in nature.

This model is represented by the DAG in Figure 5.13(b). y_{A1} and y_{A2} are PCI observations of pavements belonging to category A. x_{A1} and x_{A2} are corresponding pavement ages (or log ESAL) (the explanatory variable). α_A and β_A are the intercept and age (or traffic) effect for pavements belonging to category A. For the linear model specified here, η_{A1}, the unobserved mean of y_{A1}, is calculated as the linear combination of all the incoming nodes: $\alpha_A + \beta_A \times x_{A1}$ and $\eta_{A2} = \alpha_A + \beta_A \times x_{A2}$. α_A and β_A contribute to all data in category A. The same applies to data in category B, except that they use a separate set of parameters, α_B and β_B. As a result, data and variables for category A and category B are put into separate boxes and no link exists between them.

5.6.2 Non-spatial individual street ID based models

As each street segment may have its own characteristics, the second type of model refines the categories in the non-spatial categorical regressions by treating each street as a category itself, e.g., a separate box for each road segment as shown in Figure 5.13(c). A linear trend is fitted for each street segment. α and β are the global average intercept and the coefficient for the age or traffic effect. These two variables are shared by data in all categories (not shown but contributing to all η_{ij}s in Figure 5.13(c)). Besides, for each street segment i, ξ_i and u_i are individual street's deviations in intercept and age effect (or traffic effect in model $street - logESAL$) from the global mean. Both ξ_i and u_i are assumed to be independent and identically distributed (i.i.d.) variables (normally distributed, to be specific) and they only contribute to η_{ij} within the same box in Figure 5.13(c).

The differences between the non-spatial categorical models and the non-spatial individual models not only lie in the numbers of the categories. Furthermore, the intercepts and slopes in the non-spatial categorical models are allowed to vary without constraints, while these parameters in the non-spatial individual street ID based models have to satisfy global normal

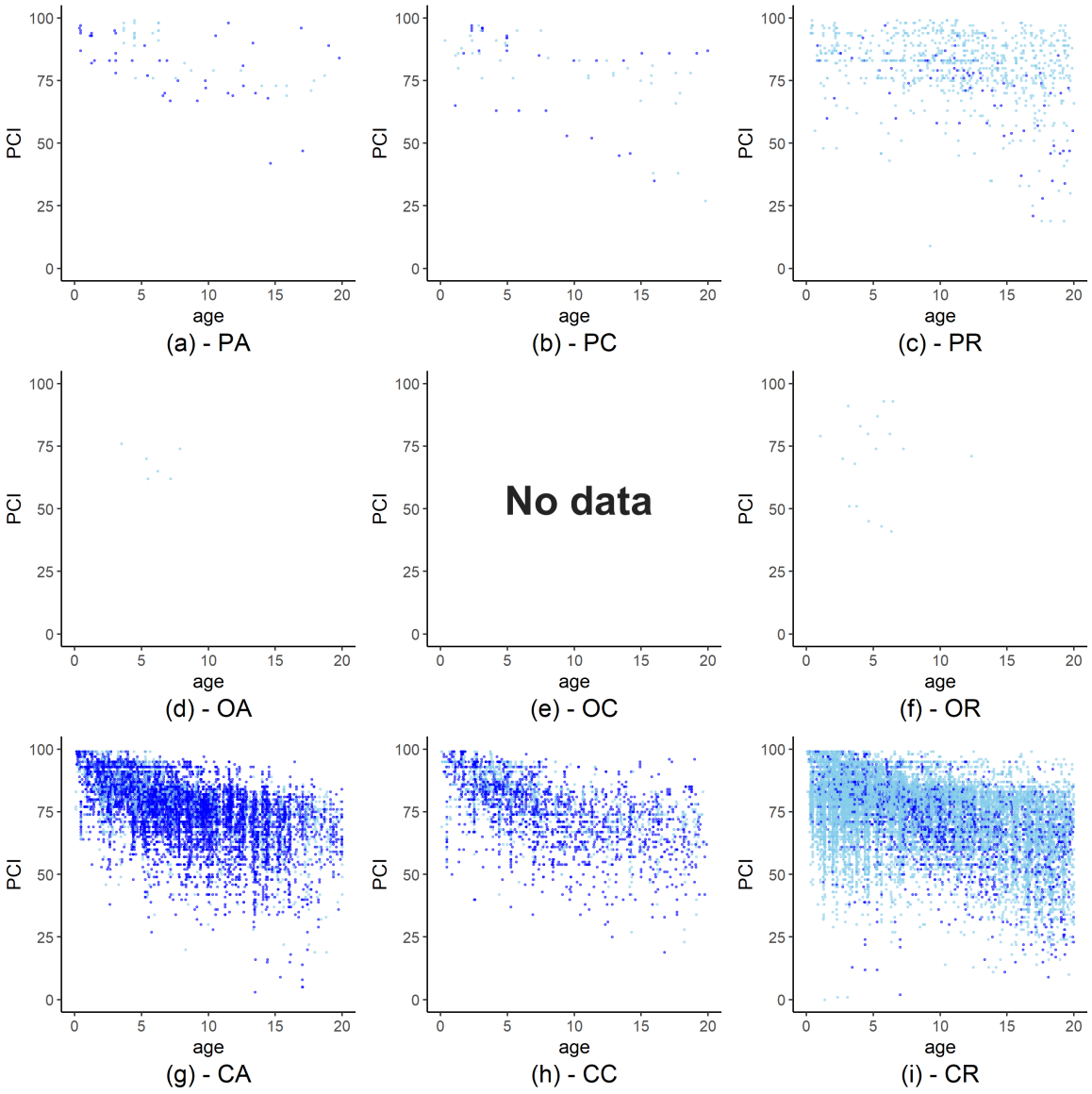


Fig. 5.14 Data categorisation for the non-spatial categorical age-based model. Light blue is for street segments without bus traffic and dark blue is for street segments with bus traffic. Meanings of other abbreviations are explained in Table 5.2.

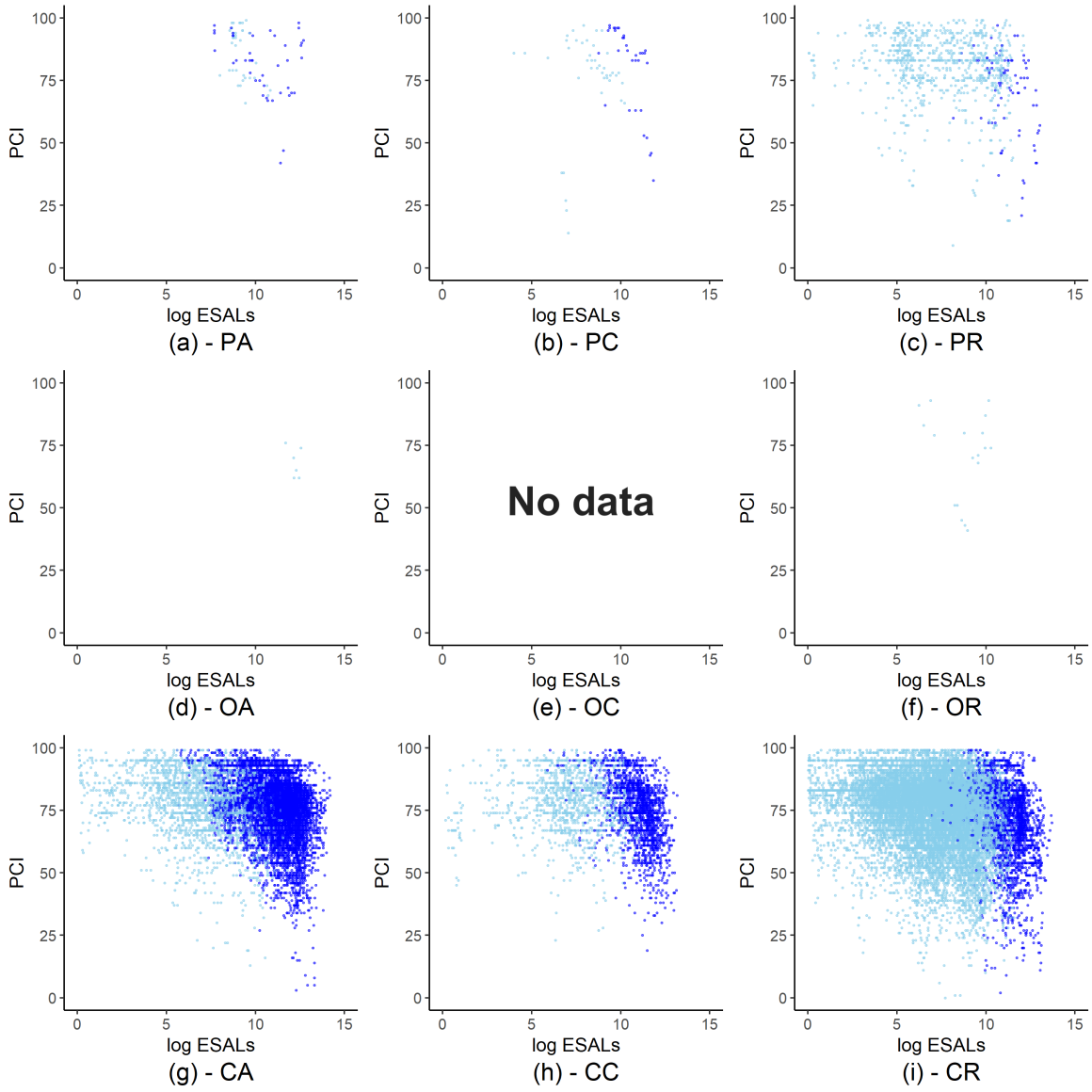


Fig. 5.15 Data categorisation for the non-spatial categorical log ESAL-based model. Light blue is for street segments without bus traffic and dark blue is for street segments with bus traffic. Meanings of other abbreviations are explained in Table 5.2.

distribution constraints, controlled by precision variables τ_u and τ_ξ. This is because, as the categories become finer, fewer data points exist in each category and regression coefficients will be significantly affected by errors in the data points. To prevent unrealistically large or small intercepts and slopes from being obtained, global constraints are thus imposed to limit their variation ranges. It will be shown in the upcoming subsection that the spatial model addresses the same issue, but with a spatially-correlated parameter.

5.6.3 Spatial models

As there are nearly 8,000 street segments in the cleaned dataset, the non-spatial individual street based models ($street - age$ or $street - logESAL$) also include thousands of parameters. However, most of the additional parameters (ξ_i and u_i) are random effects that serve to improve the model fit without revealing specific reasons or patterns. In comparison, the spatial model partitions the street-level age effect into two parts: the spatially-structured v_i and spatially-unstructured u_i. This is shown graphically in Figure 5.13(d), where η_{ij}, now having six parent nodes (α and β not shown for clarity of the Figure), is calculated as $\alpha + \xi_i + (\beta + u_i + v_i) \times x_{ij}$.

The spatially-structured v_i varies smoothly across the space and makes clusters of road segments based on their degradation rates. Specifically, v_i is assumed to follow the 'Besag' specification (Equation (5.6)), with mean value equals to the average of the spatial components of its neighbours. This is the unique feature and an advantage of the spatial model, which is to "borrow information/strength from neighbours". Neighbours are defined as adjacent road segments. The more neighbours street segment i has, the smaller the variance of v_i. Similar to the non-spatial individual street based models, the spatially-unstructured random effects (ξ_i and u_i) are also subjected to the global distribution constraints. The spatially-structured v_i, on the contrary, can be viewed as localised constraints on the parameter values.

5.7 Results

Parameter estimations for the degradation models are carried out in software R 3.4.1 [189]. Priors are specified as the following: for α, the PCI of newly constructed pavement, its physical meaning limits its value to be close to 100. A normal prior with large variance $N(0, 0.0001^{-1})$ is adopted for α. With sufficient data, it is expected that the prior is only weakly-informative, and the posterior distribution will be much narrower than the prior. The same rationale applies to β, the annual PCI degradation rate with age or traffic, whose actual

value is estimated to be a small negative number around -5 to 0, and $N(0, 0.0001^{-1})$ is again chosen as the prior. The hyperparameters τ_ξ, τ_u, τ_v and τ_g are "precision" parameters that are mathematically constrained to be positive. So the *Gamma* distribution with large variance, $Gamma(1, 0.0005)$, is used as their priors. The *Gamma* distribution is parameterised with the shape and rate parameters, while the normal distribution is parameterised with mean and variance (inverse of precision), same as the parameterisation used in Section 5.6. Note that in the previous chapter, the *Gamma* distributions are parameterised with the shape and scale parameters. While the shape and rate parameterisation is adopted in this chapter as it is a more common form in Bayesian statistics.

Resulting street-level initial conditions and degradation rates under the above specified priors are shown in Figures 5.16, 5.17 and 5.18 as maps. Moreover, Table 5.4 and 5.5 provide detailed numerical summaries of the results, including (1) values or distributional characters of the regression coefficients α, β, ξ, u and v; (2) the root-mean-square errors (RMSE) on the training and the testing datasets; and (3) the Deviance Information Criteria (DIC) for comparing Bayesian models [220], which will be explained in detail later in this section.

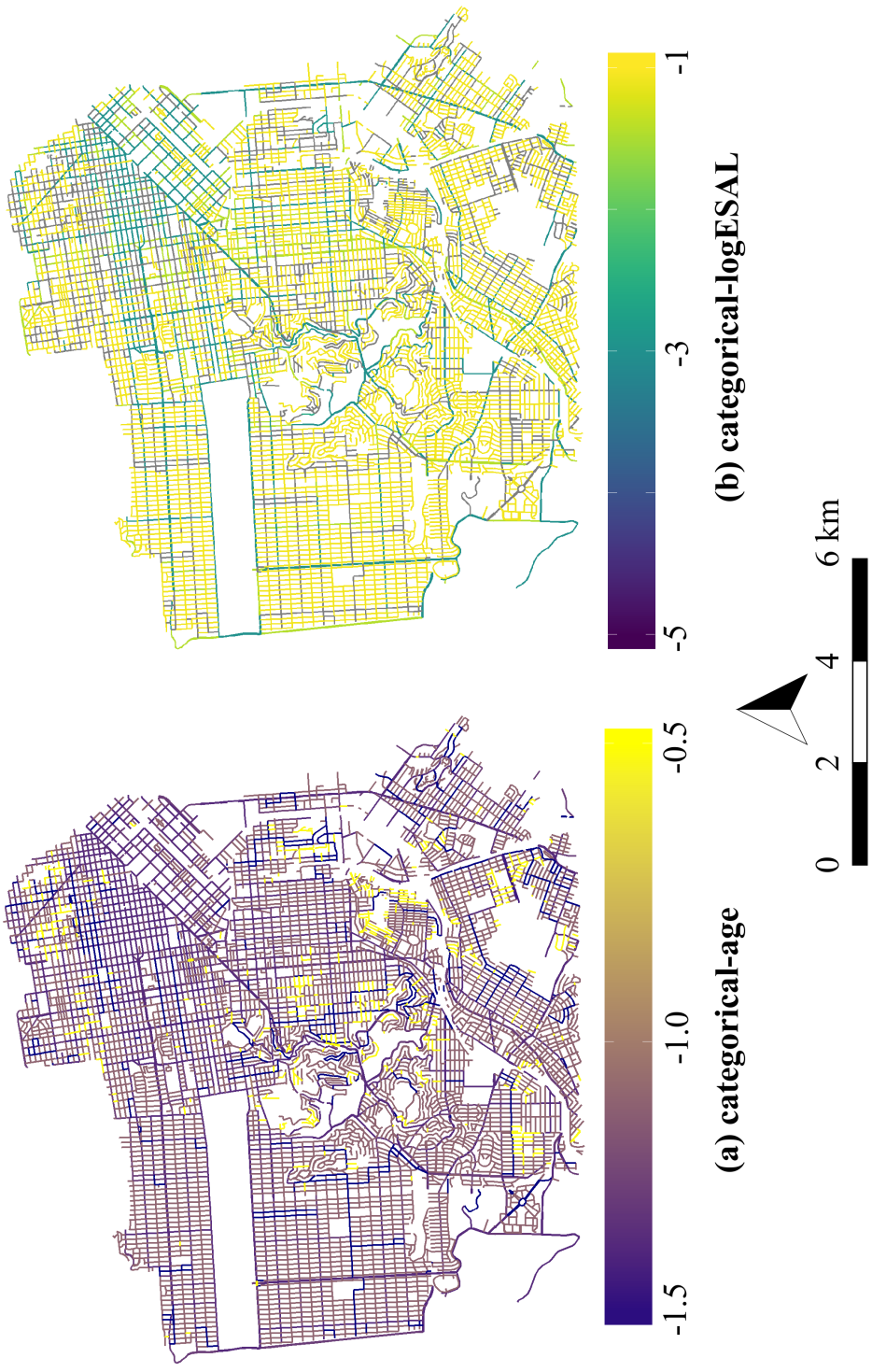


Fig. 5.16 Pavement degradation rate map based on street type categories. (a) The age model. (b) The log ESAL model.

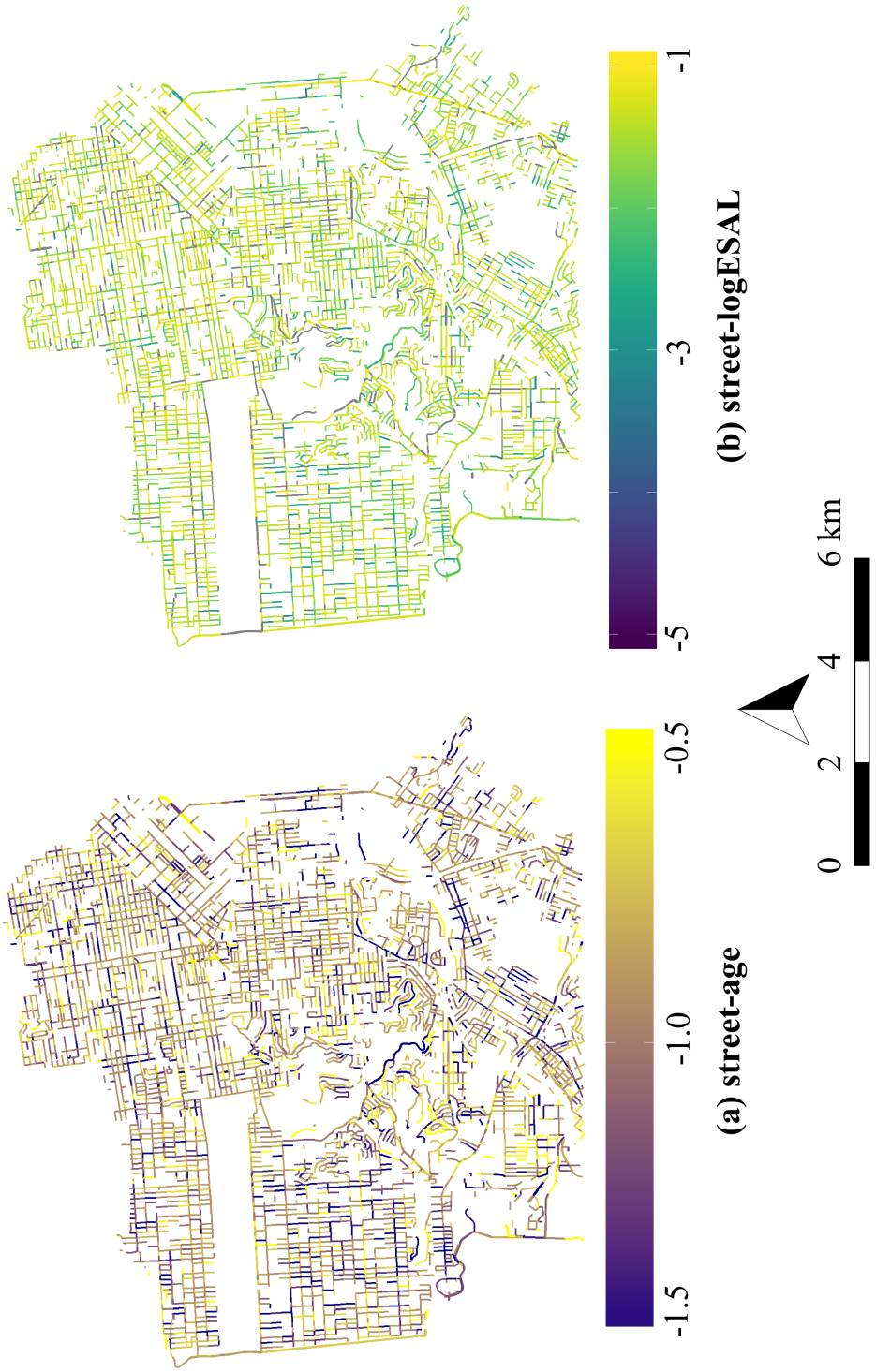


Fig. 5.17 Pavement degradation rate map based on individual streets. (a) The age model. (b) The log ESAL model.

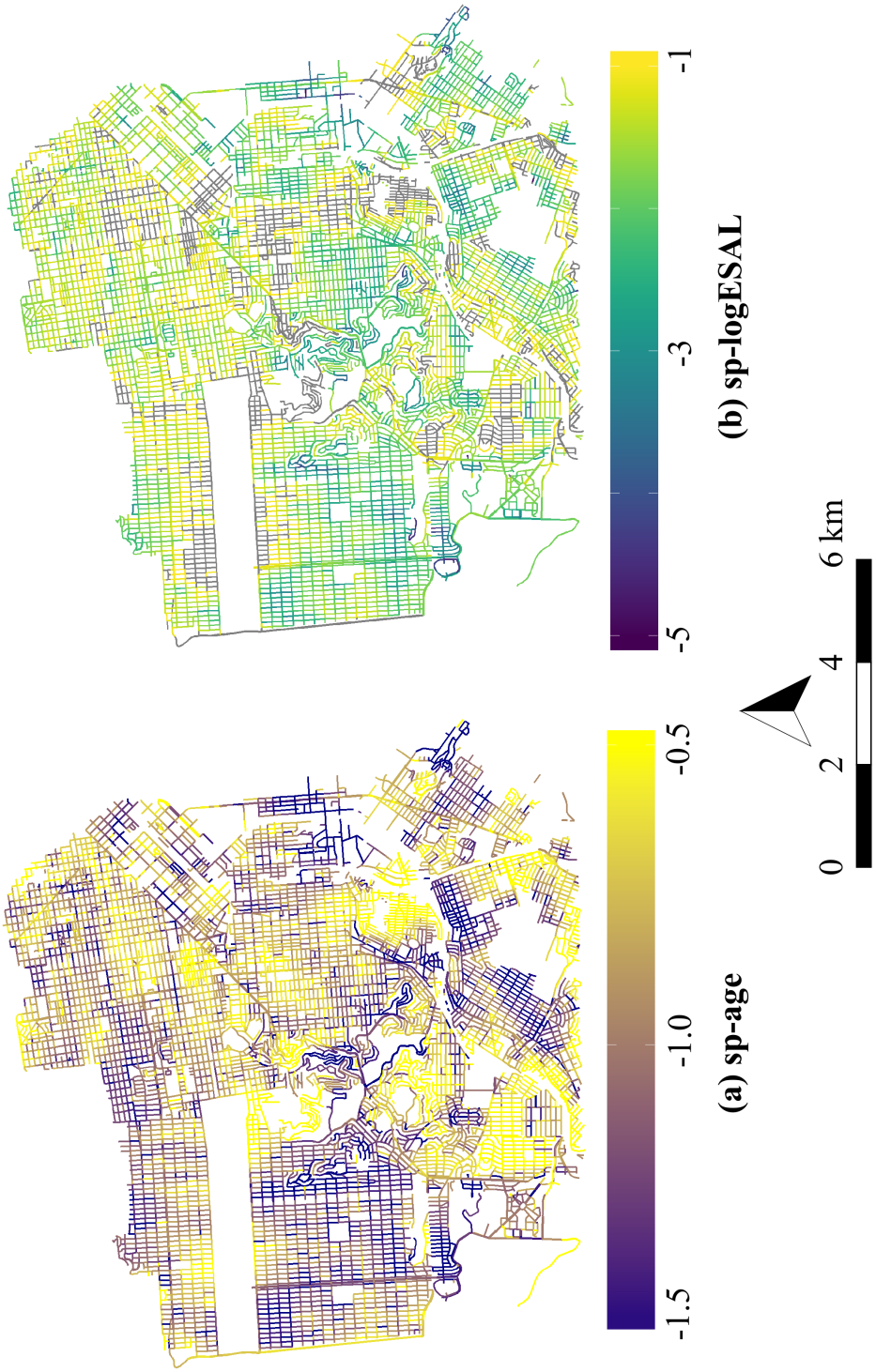


Fig. 5.18 Pavement degradation rate map of the spatial models. (a) the age model. (b) the log ESAL model.

Table 5.4 Results of the pavement age models

Model	Intercept	Slope	Performance
$categorical - age$	- non-bus routes { CA: 87.85 CC: 88.01 CR 84.44 OA: 53.93 OC, OR: no data PA: 97.33 PC: 92.05 PR 87.15 - bus routes { CA: 87.35 CC: 86.77 CR 86.99 OA, OC: no data OR: 62.95 PA: 85.01 PC: 88.30 PR 89.88	- non-bus routes { CA: -1.26 CC: -1.24 CR -1.09 OA: 2.20 OC, OR: no data PA: -1.41 PC: -1.50 PR -0.52 - bus routes { CA: -1.33 CC: -1.34 CR -1.58 OA, OC: no data OR: 1.04 PA: -0.47 PC: -0.99 PR -1.50	- Root Mean Square Error { training: 11.32 testing: 11.84 - DIC: 224504
$street - age$	- Global mean α: 85.07 - Individual deviation ξ_i { 1st quantile: -2.31 mean: 0.74 3rd quantile: 3.33	- Global mean β: -0.96 - Individual deviation u_i { 1st quantile: -0.10 mean: 0.06 3rd quantile: 0.19	- Root Mean Square Error { training: 7.05 testing: 9.14 - DIC: 209625
$sp - age$	- Global mean α: 84.98 - Individual deviation ξ_i { 1st quantile: -2.20 mean: 0.68 3rd quantile: 3.16	- Global mean β: -0.94 - spatially structured individual deviation v_i { 1st quantile: -0.20 mean: 0.02 3rd quantile: 0.23 - spatially unstructured individual deviation u_i almost 0.	- Root Mean Square Error { training: 7.09 testing: 9.10 - DIC: 209555

Table 5.5 Results of the traffic log ESAL models

Model	Intercept	Slope	Performance
categorical $-logESAL$	- non-bus routes { CA: 88.24 CC: 79.43 CR 83.53 OA: 3.55 OC, OR: no data PA: 76.91 PC: 73.93 PR 86.77 - bus routes { CA: 109.64 CC: 134.4 CR 131.19 OA, OC: no data OR: 57.61 PA: 82.69 PC: 107.70 PR 103.12	- non-bus routes { CA: -1.45 CC: -0.14 CR -1.07 OA: 5.28 OC, OR: no data PA: 1.00 PC: 0.58 PR -0.74 - bus routes { CA: -3.00 CC: -5.38 CR -5.38 OA, OC: no data OR: 1.23 PA: -0.11 PC: -2.59 PR -2.89	- Root Mean Square Error { training: 12.58 testing: 13.09 - DIC: 230602
individual $-logESAL$	- Global mean α: 89.92 - Individual deviation ξ_i { 1st quantile: -2.12 mean: 0.96 3rd quantile: 3.49	- Global mean β: -1.44 - Individual deviation u_i { 1st quantile: -0.19 mean: 0.06 3rd quantile: 0.28	- Root Mean Square Error { training: 7.93 testing: 10.06 - DIC: 215423
$sp-logESAL$	- Global mean α: 90.60 - Individual deviation ξ_i { 1st quantile: -2.00 mean: 0.75 3rd quantile: 3.10	- Global mean β: -1.54 - spatially structurediIndividual deviation v_i { 1st quantile: -0.35 mean: 0.08 3rd quantile: 0.43 - spatially unstructured individual deviation u_i almost 0.	- Root Mean Square Error { training: 7.93 testing: 9.95 - DIC: 214964

5.7.1 Non-spatial categorical models

In the non-spatial categorical models, regression coefficients (α and β) are obtained based on street type category and the resulting degradation trends are plotted alongside the data in Figure 5.19 and 5.20. For the age based model, as there are not enough data for streets with surface type 'O' (second row of the plots), their degradation trends are shown to be completely unreliable. For categories where the data points are relatively sufficient, e.g., type 'PR', 'CA', 'CC' and 'CR', the degradation trends are more sensible. Besides, for these four categories, the degradation rate for streets carrying buses (dark blue series), the slopes are all slightly larger than the streets without buses (light blue series) in the corresponding categories. However, the differences in the degradation rates are quite small, usually below 0.5 PCI per year.

For the traffic based model, as shown in Figure 5.20, the problem of the scarcity of data still prevails for most of the results in the first two rows (road surface type 'O' and 'P'). For the categories with sufficient data (type 'PR' and all the bottom row), the degradation rates of the bus-carrying streets (dark blue series) are significant faster than streets without buses (light blue series). Yet this is likely to be a modelling issue rather than the underlying behaviour of the data. For the bus-carrying streets, the cumulative traffic load (the x-axis) is significantly larger than the streets without buses, thus the data points belonging to the former all shift to the right. While the use of the log scale for the explanatory variable concentrates the data points with large traffic load horizontally, all together showing a steeper slope for the bus-carrying group. As shown by the training and testing errors of model $categorical - age$ and $categorical - logESAL$ in Table 5.4 and 5.5, the fitting of the traffic-based model (RMSE of 12.58 for the training data and 13.09 for the testing data) is worse than the age-based model (RMSE of 11.32 for the training data and 11.84 for the testing data), despite the clearer separation of data in Figure 5.20.

Additionally, the results from the non-spatial categorical models are presented in different forms as in Figure 5.16 and Figure 5.21 (leftmost clusters). They will be compared and discussed with the other two models. It should be noted that different predicting variables are used for the $categorical - age$ model and $categorical - logESAL$ model, so the coefficients of these two models (e.g., α and β) should not be compared directly.

5.7.2 Non-spatial individual street ID based models

The non-spatial individual street ID based models fit a linear trend for each street segment. The fitting and testing performance of such models are much better than the categorical models, as shown by the lower training and testing RMSEs (smaller than 8 for the training

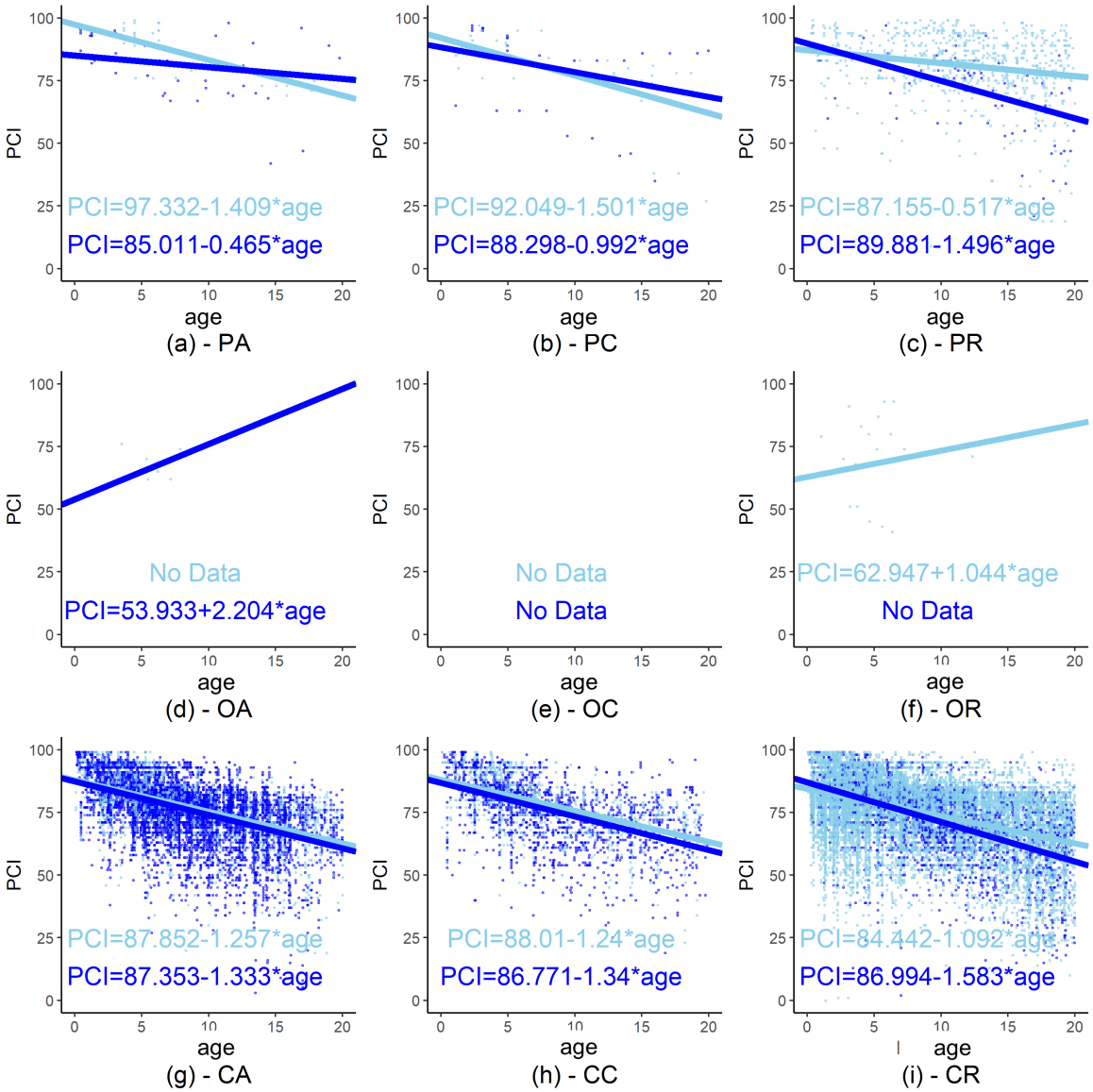


Fig. 5.19 Regression equations for the non-spatial categorical age-based model. Light blue is for street segments without bus traffic and dark blue is for street segments with bus traffic. Meanings of other abbreviations are explained in Table 5.2.

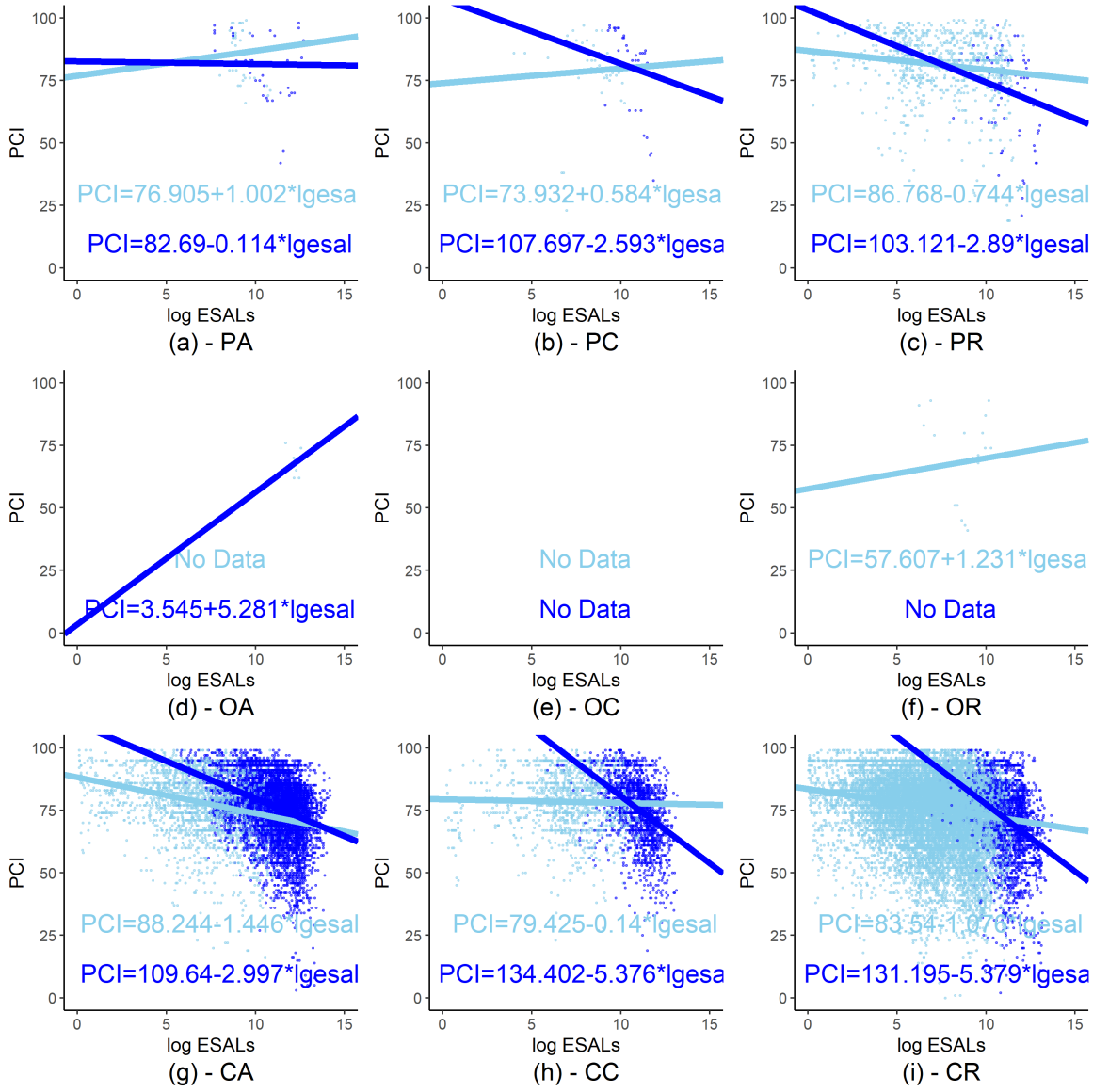


Fig. 5.20 Regression equations for the non-spatial categorical log ESAL-based model. Light blue is for street segments without bus traffic and dark blue is for street segments with bus traffic. Meanings of other abbreviations are explained in Table 5.2.

Table 5.6 Aggregating non-spatial individual street model results by surface type, functional class and bus routes

Street type functional class	Degradation rate with age		Degradation rate with log ESAL		Data size	
	bus	no bus	bus	no bus	bus	no bus
CA	-0.97	-0.89	-1.26	-1.40	2317	955
CC	-0.95	-0.90	-1.26	-1.36	699	573
CR	-1.12	-0.97	-1.47	-1.53	844	5984
PR	-0.99	-0.63	-1.50	-1.32	44	588

data and 10 for the testing data) in Table 5.4 and 5.5. Again, the age-based model performs better than the log ESAL based model. Focusing on the age-based model, the results are given at row $street - age$ in Table 5.4 suggests that on average the initial PCI condition is 85.07 and drops by 0.96 every year after.

As in this model, each street itself is one category, so factors such as surface types, functional classes and bus volumes are not modelled explicitly. Nonetheless, as one set of intercept and slope parameters has been obtained for each street, it is possible to average them by groups as shown in Table 5.6. Only groups with sufficient data points are included. It can be seen that for the age model, the streets with bus traffic degrade faster with time across all pavement surface types and functional classes, which is consistent with the results of model $categorical - age$. While for the traffic model, most of the streets with bus traffic degrade slower with increasing traffic load, opposite to the trend observed in model $categorical - logESAL$ when breaking down the categorisation to individual roads. This could be because of the influence of priors when the group size is small in the individual street based model.

The street-level initial conditions and degradation rates are shown as maps in Figure 5.17. Additionally, Figure 5.21 is provided for more visual and detailed comparison of the results. In this figure, each point stands for the regression coefficient (the age effect) for a group or a street segment. The points are clustered and coloured based on street types. Red horizontal bars mark the locations of the 10th, 20th, ..., 90th percentiles for each group. Comparing the non-spatial individual street based results (marked as $1.NSP\_ROAD\_AGE$ on the x axis) with those from the non-spatial categorical models ($0.NSP\_CATE\_AGE$ on the x axis), it can be seen that the regression coefficients are scattered even within the same street category. For example, in the age effect obtained from the non-spatial individual street based regressions, the 30th percentile is -0.36 for pavement surface type P (Portland

cement concrete), residential streets (type "PR"), while the 70th percentile degradation rate of the same type of street is -0.84, almost 2 times faster. This suggests that simple group-level results in the categorical models cannot adequately capture the individual variations in degradation trends.

5.7.3 Spatial models

The spatial models are spatial extensions of the non-spatial individual road ID based models in a way that partitions the street-specific degradation rate into two parts: the spatially structured v_i and the spatially unstructured u_i. Maps showing the resulting street-level coefficients from the spatial models are given by Figure 5.18. It can be seen that, the resulting degradation rates of the spatial model are considerably different from those of the rest in terms of the spatial pattern. For example, in the non-spatial individual street based models, degradation rates are estimated largely based on individual road segments. In comparison, the regressions of the spatial models are coordinated between several street segments within the defined neighbourhood. The degradation rate map from the non-spatial individual street based models does not exhibit any special structure or pattern (Figure 5.17), while for the spatial models, it shows regions of high and low degradation rates (Figure 5.18). This spatial pattern can be further studied to reveal underlying causes of the differences in degradation rates, such as the ground conditions, the micro-climate or the absence of sufficient drainage capacities.

To compare the spread/distributional characters of the regression results from the spatial and non-spatial models, Figure 5.21 is again used. In Figure 5.21, the resulting degradation rates from the spatial models are further divided into two parts: $SP-age(u)$ only shows the non-spatial components (i.e., $\beta+u_i$), while $SP-age(v)$ shows the spatial components (i.e., $\beta+v_i$). The variation range of v_i is significantly larger than u_i. In other words, a large part of the individual variations in degradation rates captured by the model are explained by the spatially-structured component.

5.7.4 Model comparison using RMSE and DIC

The last columns in Table 5.4 and 5.5 give metrics for model comparison. Among them, the RMSE of the training and testing datasets are used to compare the fitting and predicting abilities of the models. The training RMSE evaluates the "fitting" of the regression results compared with the data. Smaller training RMSE indicates that the model results agree well with the data. However, smaller training RMSE may also indicate that there are too many parameters used for the regression. For example, if there are 10 data points for regression,

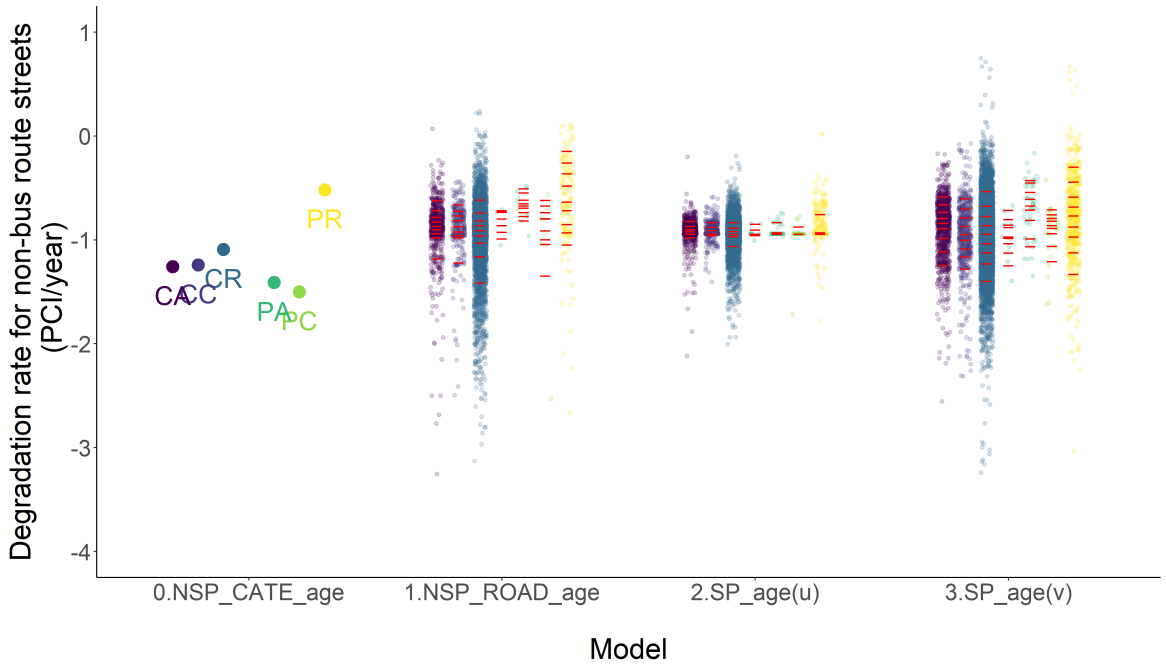


Fig. 5.21 Resulting age effects from different models based on the streets without bus traffic. See Table 5.2 for meanings of abbreviations.

a model with one parameter will probably have low training RMSE. But a model with 10 parameters will lead to a perfect fit to all the data points and lead to a training RMSE of 0. So the training RMSE alone does not adequately describe the model performance. As a result, the testing dataset is introduced to "test" the regression model on the data that have never been seen by the regression process. If the RMSE on the testing dataset is considerably larger than the training RMSE, it indicates that the model "overfits" the training data and cannot be generalised to other situations.

- The non-spatial categorical models (*categorical_age* and *categorical_logESAL*), due to their oversimplified structures, have the worst fitting/predicting performance. In fact, both models underfit the data, as the RMSEs on the testing dataset (11.84 for *categorical_age*, 13.09 for *categorical_logESAL*) are about the same as the RMSEs on the training dataset (11.32/12.58).

- The non-spatial individual street based models (*street_age* and *street_logESAL*) perform better than the non-spatial categorical models, with training RMSEs (7.05 for *street_age*, 7.93 for *street_logESAL*) and testing RMSEs (9.14/10.06) both being smaller. The RMSEs on the testing dataset are slightly higher than the training RMSE, but not large enough to be considered as overfitting.

- The spatial models have comparable training RMSEs as the non-spatial individual street based models but slightly less testing RMSEs. Specifically, the training RMSE is 7.09 when age is used as the predictor, and 7.93 for the traffic load model. The testing RMSEs are 9.10 and 9.55, respectively.

Comparatively, the improvements in fitting and predicting accuracy are not a significant advantage for the spatial models. The most important strength of the spatial models is that they have identified regions of high degradation rates. In fact, the spatial component is rather dominant in the total variability of degradation rates of the spatial models (5.21). If the pavement degradation rates do not possess spatial features, results would be expected that the values of the spatial component v_i being closer to zero for all roads. In other words, as the data size for individual pavement is small, estimating its degradation rate is greatly affected by the outliers (measurement errors). The non-spatial individual street based models address this outlier influence through identical and independent prior constraints on the parameters, while the spatial models introduce spatially correlated parameters. The spatial component is found to be dominant, also proving that the spatially correlated constraints is a reasonable problem-specific adaptation of the independent constraints. Comparing the performance of using age or traffic load as the explanatory variables side by side from Table 5.4 and Table 5.5, it is found that age always performs better as the predictor. All training and testing RMSEs of the age-based models in Table 5.4 are consistently smaller than the RMSEs of corresponding models with log ESAL as the predictor (in Table 5.5). This has important implications on the coupled pavement degradation and traffic simulation of this study and will be discussed in Section 5.8.

The DIC is another metric to compare Bayesian models. It is calculated by using either of the formula below [221]:

$$DIC = D(\overline{\theta}) + 2p_D \tag{5.12}$$

$$DIC = \overline{D} + p_D \tag{5.13}$$

where D (deviance) is a goodness-of-fit statistics equalling to $-2log(likelihood)$. Smaller deviance indicates a better fit to the data. $D(\overline{\theta})$ is the deviance at posterior means (a classical point-wise measure of model fit) and $\overline{D}$ is the posterior mean deviance (a Bayesian measure of model fit). $p_D = \overline{D} - D(\overline{\theta})$ is the effective number of parameters and reflects the complexity/degree of "overfitting" of the posterior distributions. p_D is included in DIC to penalise complex models. DIC combines the goodness-of-fit measure (D) and the effective number of parameters (p_D), thus reflecting both model fit and complexity. Smaller DIC is better as it indicates a more desirable balance of fit and complexity. The absolute value of

DIC is not meaningful and models are compared based on their differences in DIC. From Table 5.4 and 5.5, the spatial models have the smaller DIC than non-spatial models and the age-based models also have smaller DIC compared to the log ESAL based models.

5.7.5 Prior sensitivity analysis

The sensitivity of Bayesian results on the choice of prior distributions has been discussed extensively in many previous studies [6, 90]. For multilevel or other complex models, even the flat, uninformative uniform prior distributions can become restrictive and informative on the transformed parameters. Thus, it is important to test a diverse range of prior distributions and to see whether the same conclusions can be reached under different prior choices. In the Bayesian models of this chapter, prior distributions are specified for the following model parameters: the global intercept α, the global degradation rate β, the global precision τ_g and precisions for street-level coefficients τ_ξ, τ_u and τ_v. As α, β and τ_g are global variables, there should be sufficient amounts of data to lead to accurate posterior estimations.

However, the global degradation rate β is an important parameter for the pavement degradation analysis. The regression results for β based on the default prior distributions are rather small. For example, for the age-based spatial model, the mean of β's posterior distribution is less than 1. This means that on average, the pavement condition drops by only 1 PCI per year. Consider the time takes for the PCI to decrease from 100 to 50 as the pavement life, results of $\beta = 1$ indicates that the pavement life would be 50 years. This is considerably longer than the generally perceived life of the pavement (about 15 to 20 years). As a result, it is crucial to check if the small β is indeed embedded in the data (due to data collection errors, data cleaning procedures, etc.) or due to the modelling assumptions (e.g., prior influence). The default prior for β is a normal distribution $N(0, 0.0001^{-1})$. Two other prior distributions are tested, including $N(-100, 0.0001^{-1})$ and the narrow $N(-100, 0.01^{-1})$, as listed in Table 5.7.

Besides, for the street-level parameters τ_ξ, τ_u and τ_v, their prior sensitivities are also tested as there are relatively low numbers of data points for each street. *Gamma* distribution is the usual choice of prior for precision parameters. Four *Gamma* priors with different distributional characters (labelled D, E and F in Table 5.7) are tested first. It is possible to use other probability distributions as priors. However, as some streets only have two to three condition observations in the study period, uniform prior becomes too uninformative that the INLA process fails to complete [77]. As a substitute, normal priors with large variances (labelled G, H and I) are tested instead.

Table 5.7 Prior combinations in the sensitivity test

Label	Prior[1]	α	β	Training RMSE	Testing RMSE	DIC
A (Default)	$\beta \sim N(0, 0.0001^{-1})$ $\tau_\xi, \tau_u, \tau_v \sim Gamma(1, 0.0005)$	mean=84.9818 sdev=0.1100	mean=-0.9418 sdev=0.0177	7.09	9.10	209555
B	$\beta \sim N(-100, 0.0001^{-1})$	mean=84.9775 sdev=0.1098	mean=-0.9410 sdev=0.0177	7.10	9.10	209548
C	$\beta \sim N(-100, 0.01^{-1})$	mean=84.9850 sdev=0.1097	mean=-0.9424 sdev=0.0177	7.10	9.10	209558
D	$\tau_\xi, \tau_u, \tau_v \sim Gamma(0.1, 0.1)$	mean=84.9782 sdev=0.1100	mean=-0.9414 sdev=0.0177	7.10	9.10	209515
E	$\tau_\xi, \tau_u, \tau_v \sim Gamma(0.01, 0.01)$	mean=84.9801 sdev=0.1098	mean=-0.9413 sdev=0.0177	7.09	9.10	209546
F	$\tau_\xi, \tau_u, \tau_v \sim Gamma(0.001, 0.001)$	mean=84.9819 sdev=0.1100	mean=-0.9418 sdev=0.0177	7.10	9.10	209564
G[2]	$\sigma_\xi, \sigma_u, \sigma_v \sim N(0, 0.001^{-1})$	mean=84.9637 sdev=0.1085	mean=-0.9361 sdev=0.0175	7.06	9.09	209455
H	$\sigma_\xi, \sigma_u, \sigma_v \sim N(0, 0.00001^{-1})$	mean=84.9774 sdev=0.1100	mean=-0.9407 sdev=0.01758	7.09	9.10	209537
I	$\sigma_\xi, \sigma_u, \sigma_v \sim N(100, 0.00001^{-1})$	mean=84.9775 sdev=0.1101	mean=-0.9406 sdev=0.01778	7.09	9.10	209532

[1] $N(a, b^{-1})$ is the normal distribution parametrised by mean a and precision b^{-1}. $Gamma(c, d)$ is the *Gamma* distribution parameterised by shape c and rate d.

[2] Here the standard deviation parameter $\sigma_\xi = 1/sqrt(\tau_\xi)$. It follows normal distribution but truncated to positive [111].

Based on the results of the estimated parameters, RMSEs and DIC under different prior specifications, it can be seen that the outcomes of the regressions do not vary greatly given different priors. Specifically, the degradation rate parameter β is always around -0.94, indicating that slow degradation rate obtained is likely to be caused by the input data or the data cleaning procedure, rather than the modelling process. One most likely reason behind this small β is as the following: in the SF PCI observation records, minor maintenance events are not documented. However, these minor maintenance (e.g., crack sealing) does have an effect on improving the pavement conditions and preventing further degradation, which will make the pavement condition degradation appears to be smaller over time. In the future, this problem can be addressed by, e.g., more frequent measurements through advanced sensing techniques.

5.8 Discussions

In this section, two topics related to the modelling of the pavement degradation in SF are discussed. The first is a reflection on the (lack of) interactions between the pavement degradation and traffic distribution. The second part comments on the suitability of the spatial model for the simulations of the next chapter on transport system-wide CO_2 emissions. Although it is acknowledged that the outcomes of the pavement degradation model are not completely satisfying, due to the relatively large RMSEs, it still provides a situation or case study specific knowledge that is better than generic models borrowed directly from elsewhere.

5.8.1 Interaction between traffic and pavement degradation

Despite the efforts in this chapter to include realistic spatio-temporally varying traffic data in the pavement degradation analysis, the traffic-based models are found to be consistently underperforming compared to age-based models with similar structures. There are many reasons behind this.

First of all, the traffic data supplied into the pavement degradation models are simulated data rather than actual traffic counts. Although the traffic simulation model has been validated preliminarily with an official traffic simulation model from the Bay Area, there is still no guarantee that the simulated traffic volume (and the ESAL) reflect the real load history that each pavement segments have experienced. At the beginning of this study, it has been attempted to obtain the actual historical traffic volumes from open data sources or local agencies, however unsuccessful. The annual traffic counts data obtained from the San Francisco County Transportation Authority (SFCTA) are based on manual counts on a

few specific days and the files are not digitised. So traffic simulation data from the traffic simulations appear to be the only available choice.

Secondly, the pavement condition data based on manual surveys inevitably contain errors. According to Tan and Cheng [225], the variability in the PCI measurements can be as high as 7 to 10 PCI points. As sometimes there are only 3 to 4 data points per street for building the regression model, a deviation of 10 PCI can severely bias the estimation of the degradation trend. Besides, the small maintenances such as pothole patching and crack sealing are not documented. While these small repair works may improve the pavement condition locally and prevent further degradations. It is likely that for two streets having comparable traffic load, one has received a few more small maintenances over the past 20 years. This lack of small maintenance records could mean that, even with the actual historical traffic data, it may not be helpful in terms of building a more precise model.

Thirdly, as shown in the scatter plot of PCI and cumulative ESAL without the log transformation (Figure 5.11(b)), hardly any downward degradation trend is available even though the cumulative ESAL without the log transformation was initially devised to evaluate the damage potential of diverse vehicle load conditions. The log transformed ESAL being used is effectively converting the relationship back to age:

$$PCI_{ik} = \alpha_i + \beta_i \times log(ESAL_{ik}) \tag{5.14}$$

which is approximately equivalent to

$$PCI_{ik} = \alpha_i + \beta_i \times log(age_{ik} \times traffic_{ik}) \tag{5.15}$$

$$PCI_{ik} = \alpha_i + \beta_i \times log(traffic_{ik}) + \beta_i \times log(age_{ik}) \tag{5.16}$$

where PCI_{ik}, $ESAL_{ik}$ and age_{ik} are the PCI, ESAL and age of pavement segment i in year k, respectively. The above equations show that, after the log transformation, the traffic is now affecting the initial conditions, which is not the dynamic effect between traffic load and pavement degradation that the analysis is seeking. Thus any effect of the log ESAL based models is still largely reflecting the age effect. As neither of the ESAL and log ESAL used in this analysis turns out to be a satisfying predictor, it has to be concluded that age should be the preferred choice of the explanatory variable to predict the pavement degradation in this specific study.

5.8.2 Suitability of the spatial model for further analysis

A weakness in this study, as shown in the Results section (Section 5.7), is that for all the models presented, their training and testing errors are around 7 to 10 PCI. This is not surprising given the wide scattering of the input data. Actually, according to Tan and Cheng [225], the variabilities in the PCI raw data from manual surveys are about the same level. A limitation in the study is that degradations are only modelled linearly with age. Linear models are adopted because they show the key trends more directly than other complex models. But the downside is that the model fit would be compromised. The large RMSEs in all three models are likely due to that their simple forms are not fitting the errors in the measurements.

However, on the other hand, the spatial models indeed produce a better fitting and testing performance than the categorical models, which is typically used for pavement modelling. Also, the spatial models are able to adjust to missing data. This is demonstrated in Figures 5.17 and 5.18. Comparing the non-spatial individual street based models (Figures 5.17) and the spatial models (Figures 5.18), it can be seen that the former fails to generate regression coefficients for streets without any observations (marked by the small empty locations in Figure 5.17). While for spatial models, due to the borrowing of information from nearby roads, regression coefficients can be obtained for all the streets in the network. Further more, the spatial model is compensating for the absence of important factors by the assumption that the streets neighbouring each other are more likely to share similar materials, construction qualities, ground conditions, micro-climates and so on. As a result, despite the limitations in the degradation study, the spatial model (age-based) is still the best choice among all alternatives to be used for further analysis.

For this case study of modelling pavement degradation in SF, the spatial model only wins by a slender lead in terms of accuracy. But the real advantages of the spatial model are within the analysis: first of all, it is able to estimate the degradation parameters for road sections with missing or erroneous observations by borrowing information from adjacent sections, while the information in the spatial structure of the pavement network is lost in non-spatial models. Moreover, it can visually illustrate regions where pavements degrade faster than average. These regions do not necessarily have the worst pavement conditions, but they may need maintenance more often in the long term. Local engineers can be consulted, or site investigations conducted, as for the underlying causes. The latter makes spatial models particularly useful in the real practice, as it can assist asset managers to narrow down their attention to a smaller region.

Chapter 6

Integrating traffic and pavement modules for emission mitigation simulations

6.1 Introduction

The previous chapters focus on developing tools, including a mesoscopic traffic mobility model and a spatial-explicit pavement degradation model, both at city-scale with key inputs from real observations from San Francisco (SF). In this chapter, the two models are integrated together into a transportation system model. The integrated mobility and infrastructure model will be used to study the effects of different CO_2 mitigation measures, including the eco-friendly route choice (eco-routing) and the emission reduction oriented pavement maintenance (eco-maintenance). The focus is to bring the carbon mitigation strategies from two distinct perspectives (traffic operation and infrastructure asset management) into the same evaluation framework, to compare the characteristics of each strategy and to investigate the strength of interactions when implementing multiple strategies together.

The transportation sector is the largest source of greenhouse gas (GHG) emissions in California. It is estimated to account for 39% of the total inventory in 2016, among which the passenger vehicles are the major contributor [36]. Many mitigation measures have been proposed to address the GHG emissions in transportation. From the pavement asset management perspective, this includes the adoption of new materials or better maintenance scheduling [193]. From the traffic operation side, emission mitigation opportunities have been identified such as fuel economy improvements, congestion management, freight logistics optimisation, car sharing and the recent adoption electric and autonomous vehicles [97].

However, due to the complex interactions between various CO_2 mitigation measures as well as the knowledge barrier across fields, it is still unclear, e.g., how costly and effective the CO_2 mitigation strategies based on pavement asset management would be compared with the alternatives from the traffic control perspective, and whether the interactions between different strategies would complement or impair the outcomes of one specific measure.

This chapter aims to answer such questions with an integrated city-scale traffic mobility and pavement degradation model. The idea behind integrating the two models can be thought as the following: the mesoscopic traffic simulation can model scenarios regarding bottom-up behaviour changes from individual drivers, such as switching to lower emission routes. While the pavement degradation model represents the infrastructure condition change and the related asset management process, reflecting top-down approaches from a government agency to reach the goal of carbon emission reduction.

The traffic movements and pavement conditions are intertwined in many aspects. For example, degraded pavements lead to an increase in rolling resistance and fuel consumption, which may affect the route choices under an eco-routing scheme. Also, the uneven distribution of vehicles over the network at different hours of a day can affect pavement asset management decisions as for when and where to carry out roadworks. Integrating two city-scale models has the benefit of ensuring complex scenarios as the above to be included in the analysis. In this chapter, two scenarios will be investigated as for their effects on the efficiency and carbon emissions of the traffic system: (1) eco-routing: the dynamic change of driver behaviours to adapt to the traffic conditions as well as the pavement condition reductions and improvements; (2) eco-maintenance: the adaptation of pavement maintenance scheduling (site selection) according to the evolving traffic flow patterns in (1), so as to repair the most critical roads (offering the highest return on CO_2 reduction). As noted in the methodology review in Chapter 2 and summarised in Table 6.1, the adopted traffic mobility model and pavement degradation model are not the most advanced or sophisticated tools that are available in the field of traffic simulation and pavement degradation modelling. However, they are able to capture key variations (route choice, maintenance site selection) to answer the research question and are found to be feasible given the availability of public data and the time scope of this research. The innovation here is to combine the existing traffic simulation and pavement degradation modelling methodologies together to form a system-level understanding of the performance of the transportation system.

Table 6.1 Strength and limitations of the adopted traffic mobility and pavement degradation models

Model	Strength	Limitations
Traffic model	• Detailed representation of all drivable streets in SF as a directed graph, including speed limit, capacity and slope for each road link. • Hourly travel demand based on TNCs (Uber/Lyft) data. The percentages of TNC trips among all vehicular trips are sourced from calibrated models. • Route choice of individual travellers. Graph weights updated frequently (20 times in an hourly time step). • Output hourly link-level speed and traffic volumes, which can be linked with existing research to estimate traffic-related emissions. • City-scale coverage and link-level resolution matching the pavement condition observation data, allowing the two models to be integrated at the same scale.	• No sub-link vehicle dynamics (accelerations, stops at traffic signals, etc.) • Link-level speed calculated from the simplified BPR curve with simple capacity calculation procedure. • Not activity-based, not including complex socio-demographic or individual characteristics and no feedback update of travel demand (e.g., peak avoidance behaviours). • Not including buses, trucks or electric vehicles. • Despite the efficiency , the shortest path searching part is still the computational bottleneck, preventing the detailed day-by-day simulation over a decade-long analysis.
Pavement degradation model	• Individualised degradation trend for each street in SF. • Found to have better fitting and testing performance based on the observation data compared to the conventional categorical models. • Spatial correlations are used to infer the degradation rates for streets without sufficient observations.	• Data for building the pavement degradation model highly scattered. Even a large number of variables (individual street based model) cannot capture the variations. • Key inputs, such as the construction date, quality or localised maintenance activities, are missing. • Traffic load are found to be not a good predictor for pavement degradation in this study, thus limiting the interactions between the traffic model and the pavement model.

The pros and cons of the adopted traffic and pavement models are given in Table 6.1. For the traffic model, it is basically a static traffic assignment tool with hourly travel demand disaggregated to the individual traveller level. It has high spatial resolution (e.g., containing all the drivable roads in SF, the case study city), but not so detailed as to include the changes of geometry, traffic lights, etc. The temporal resolution of the traffic model is given by the time step length, which is one hour for the case study. The travel demand is assumed to be fixed, thus no activity rescheduling is involved. However, individual trips are explicitly modelled and each traveller can choose the respectively optimum path, which involves heavy path finding calculations running on the High Performance Computers (HPC). Also, the multi-modal traffic network is not considered in the current version of the model. The adopted traffic model is efficient due to the static assignment procedure, the trip-based fixed demand and the use of the macroscopic volume-delay relationship. In addition, it models the route choice with great diligence by calculating the optimum path for each traveller, unlike to pre-calculate a few route choice sets as used in existing studies. It is believed that the effort spent in calculating the routes for individual travellers are important for the simulation-based CO_2 emission quantification, as the key promise of the eco-routing strategy is related to the flexible route selection of travellers under various traffic and pavement degradation scenarios.

For the age-based spatial pavement degradation model, it is adopted as it can provide an individualised degradation trend for each street segment in the study area. Compared with the categorical models that fit a degradation curve for each type of pavements, the spatial model has better accuracy in matching the observation data. However, a limitation of the pavement degradation model comes from the wide scattering in the observation data, which is possibly due to survey bias and is found to be hard to eliminate. Also, traffic load is not fed back to the pavement degradation as the age-based model is found to have better accuracy.

The other important consideration when developing these models is the matching of the scale. Currently, both the traffic and the pavement models operate at the city-scale and have link or street-level spatial resolutions. The pavement degradation is a long-term process, and the traffic model can be run rather efficiently so as to allow the temporal coupling with the pavement degradation model. Due to these reasons, the two models are adopted for the simulation of the low carbon strategy scenarios.

6.2 Methodology

To achieve the goal of simulating emission mitigation scenarios involving both traffic operation and pavement asset management strategies, the methodology framework shown in Figure 6.1 is used to integrate the traffic mobility simulation model and the infrastructure

degradation model. The output network-wide evaluation metrics include the total CO_2 emissions, vehicle hours travelled (VHT), vehicle kilometres travelled (VKMT) and the average Pavement Condition Index (PCI). The framework involves four steps.

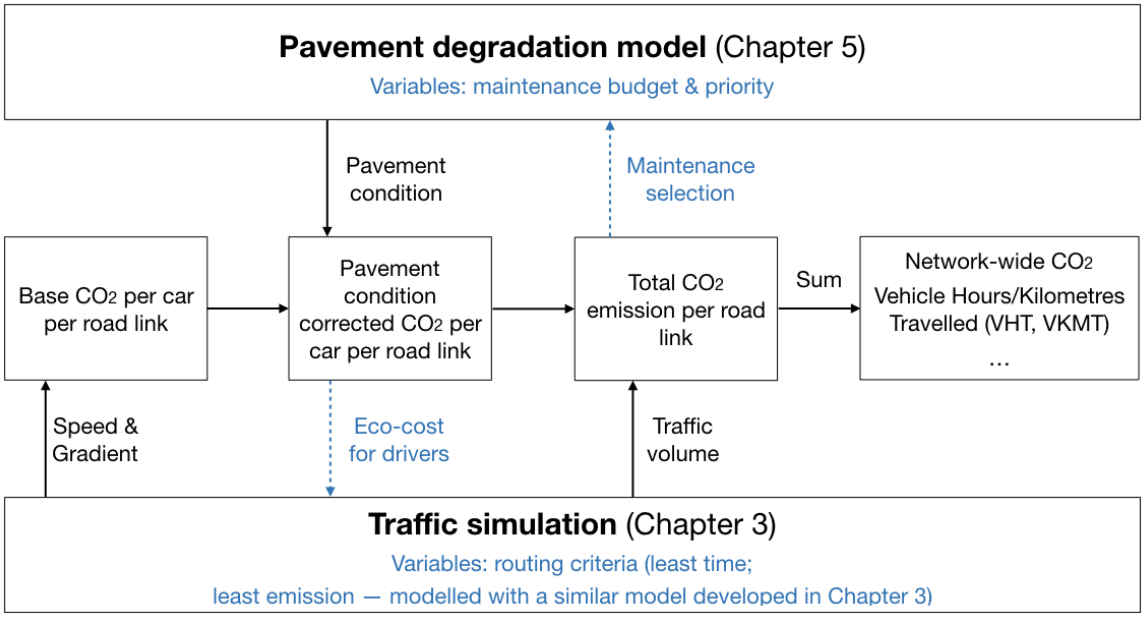


Fig. 6.1 Framework of integrating city-scale traffic model and pavement degradation model for eco-driven maintenance.

Step 1. Base CO_2 emission factor calculation

Base CO_2 emission factor refers to the CO_2 emitted by a single vehicle traversing the road link at a given speed, not considering the impact of pavement condition. The widely cited formula by Barth and Boriboonsomsin [15] is adopted for its simplicity (Equation 6.1). The regression relationship of Equation 6.1 is based on GPS trajectories of probe vehicles running on Southern California freeways (data on local roads are also collected but their emissions are not analysed). The trajectories are first matched to loop detectors in the freeways and then cut into snippets with the same level of services (LOS, a congestion measure). The trajectory snippets are next fed into the Comprehensive Modal Emissions Model (CMEM) [210] to obtain the unit distance carbon emissions. The average speed of each trajectory snippet is then obtained and plotted against the unit carbon emissions (Figure 6.2). Although the scatter plot is made with data obtained from the freeways, they are obtained under a wide variety of LOS, from LOS A (best) to LOS F (worst, congested), which justifies the applicability of the

relationship to urban roads.

$$ln(y_{CO_2})\ (g/mile) = b_0 + b_1 v + b_2 v^2 + b_3 v^3 + b_4 v^4 \quad v \text{ in } mph \tag{6.1}$$

where y_{CO_2} is the CO_2 emission in $g/mile$, v is the average trip speed in mph. $b_0 \sim b_4$ are fitted coefficients (no physical meaning). The authors provide two sets of coefficients as shown in Table 6.2: the steady-state coefficients for vehicles running at constant speeds, and the real-world coefficients that consider the higher CO_2 emissions associated with the stop-and-go driving [15]. It should be noted that the speed variable v (x-axis of Figure 6.2) refers to the average speed in a trajectory snippet, including the effects of accelerations and decelerations. For the relatively low average speed regime (e.g., $v < 55\ mph$), the congestions are getting worse and the effects of accelerations and decelerations are more prominent. This is illustrated by the increasing gap between the "real-word activity" curve and the "steady-state activity" curve in Figure 6.2, as the latter shows the CO_2 emissions to be expected for driving at a constant speed.

As the traffic simulation module outputs the link-level average speed under congested conditions, the real-world coefficients in Table 6.2 are more suitable and are adopted for the calculation in this study. The imperial units in this as well as subsequent formulae are converted to metric units for calculation and results presentations.

By using Equation 6.1, vehicle emissions can be estimated directly from the aggregated link-level speed, rather than using the detailed second-by-second trajectories. Some may argue that the link-level average speed alone cannot adequately capture the vehicle dynamics such as the accelerations and decelerations [2]. However, obtaining detailed sub-link vehicle motions requires either computationally expensive microscopic simulations (even the queue-based MATSim is not enough as it does not give results about sub-link dynamics such as accelerations or second-by-second trajectories) or GPS trajectories from a small number of probe vehicles. This limits the scaling of the application to a large network. While information of the aggregated link-level speed is easier to obtain both computationally and in reality. For example, some navigation services such as Google Maps and TomTom provide historical or real-time link-level traffic speed information (at a cost), but not the instantaneous vehicle speed and acceleration. A similar approach has been adopted in Zeng et al. [252], where link-level aggregated quantities are used to calculate emissions.

The terrain in SF is hilly and road gradients have been demonstrated in many studies to have a great impact on fuel consumption and fuel emissions [28, 56, 85]. Generally studies found about 100% increase in fuel consumption when road gradient increases from flat to 5%. As it is unclear how these results extrapolate further beyond, i.e., to the SF streets with maximum gradient of 30%, a staged linear relationship between road grade and fuel

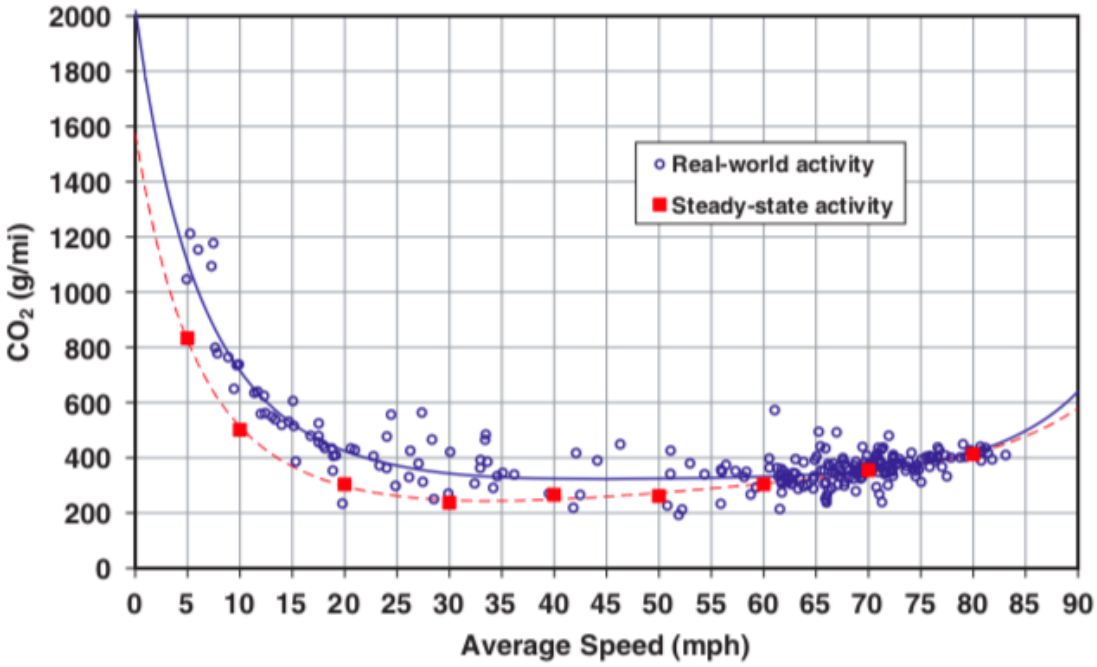


Fig. 6.2 CO_2 emissions as a function of average trip speed (Figure 3 in Barth and Boriboonsomsin [15]).

Table 6.2 Derived parameters for Equation 6.1 (Table 1 in Barth and Boriboonsomsin [15]). The coefficients are obtained from curve fitting and have no physical meaning. They are designed to be used with the variables in imperial units in Equation 6.1.

	Real-World	Steady-State
Sample size	241	9
R^2	0.668	0.992
b_0	7.613534994965560	7.362867270508520
b_1	-0.138565467462594	-0.149814315838651
b_2	0.003915102063854	0.004214810510200
b_3	-0.000049451361017	-0.000049253951464
b_4	0.000000238630156	0.000000217166574

consumption is assumed:

$$y_{CO_2,slope} = \begin{cases} 0.2 * y_{CO_2,level}, & gradient \leq -0.05 \\ (1 + 0.16 * \%gradient) * y_{CO_2,level}, & -0.05 \leq gradient \leq 0.15 \\ 3.4 * y_{CO_2,level}, & gradient \geq 0.15 \end{cases} \quad (6.2)$$

where $y_{CO_2,level}$, $y_{CO_2,slope}$ are the emissions at a levelled ground and on a sloped road. $\%gradient$ is the road gradient in percent. When the road gradient is above 15%, Equation 6.2 predicts that the unit emission is 3.4 times that on a flat terrain. When driving down a slope with gradient of -5% or below, the unit emission is only one fifth of the value on a flat surface. This curve follows the road gradient and fuel economy relationship in Boriboonsomsin and Barth [28] and is also in agreement with the conventional fuel vehicle fuel economy line chart in United States Department of Energy [232]. For roads with gradient between -5% and 15%, the emission correction factor is linearly interpolated.

Step 2. Calculating CO_2 emission factor considering pavement degradation

The emission data used to build Equation 6.1 are obtained from models based on driving pattern simulations inside laboratory with dynamometer [210]. Thus, it is assumed that the emissions calculated by Equation 6.1 do not reflect the influence of pavement roughness on the real roads. As a result, a pavement roughness correction factor is adopted to reflect the impacts of pavement roughness on CO_2 emissions. Pavement conditions that affect CO_2 emissions usually include the macroscopic texture (indicated by Mean Profile Depth, MPD), roughness (indicated by International Roughness Index, IRI) and deflection. Among them, deflection impact can be ignored for passenger cars as they are not heavy enough to generate deflections on the pavement surface [238]. However, as there is no information on the MPD or IRI of the local streets in case study area [66], these indicators need to be calculated based on existing information i.e., the PCI.

Both PCI and IRI are pavement condition indicators. As introduced in Chapter 5, PCI was developed by the US Army Corps of Engineers and is used as a comprehensive index to evaluate pavement surface distresses (ravelling, rutting, cracking, etc.). It is on a numerical scale from 0 to 100, with 100 being the best condition. On the other hand, the IRI indicator was proposed by the World Bank in the 1980s and measures the cumulative suspension motion in a moving vehicle over travelled distance [179]. An IRI of 0 m/km indicates perfectly smooth road surface. Highways in California usually have IRI between 1-5 m/km [241]. For less important roads, the IRI can be higher (more degraded). Multiple relationships have been proposed in the previous research of IRI and PCI. Park et al. [179] proposed a regression model based on 62 pavement sections in North America and Canada, but suggesting that the obtained relationship is only suitable for low IRI roads due to the limitation of the data behind the regression model. Arhin et al. [10] proposed several linear relationships between IRI and PCI using data from a dense urban area, including the relationship for composite pavements (the category that most pavements in SF belong to) with 167 observations. Furthermore,

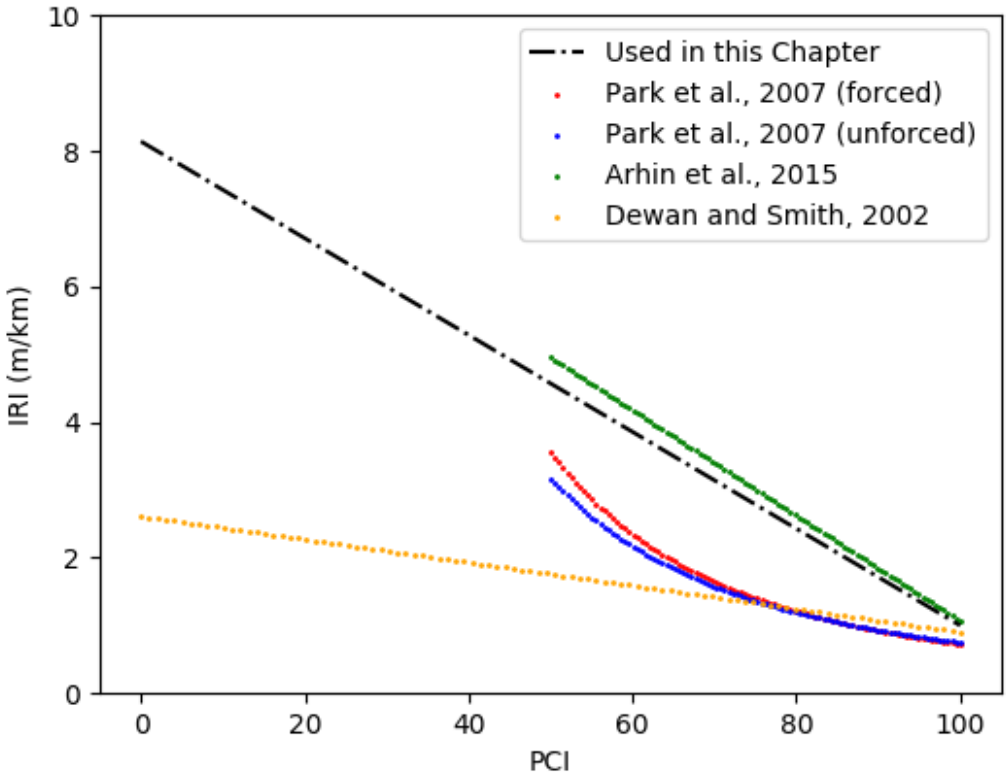


Fig. 6.3 Relationships between PCI and IRI proposed in previous research.

another study utilised specifically 39 data points from states highways in the San Francisco Bay Area [66]. These relationships from existing literature are plotted together in Figure 6.3.

As the relationships proposed by Park et al. [179] (red and blue curve) and Arhin et al. [10] (green line) are obtained only with observations of pavements at relatively good conditions, only the part of the curves with PCI larger than 50 are shown. The relationship by Dewan and Smith [66] instead covers the whole range of possible PCI values. However, since it is mainly developed from highway data, the predicted IRI < 3 m/km while PCI is 0 might be low for city streets. In the end, a straight line that is roughly in the middle of these proposed relationship is adopted, as shown by the black dot dashed line in Figure 6.3 and Equation 6.3. In this study, the degradation of the local streets are studied. As a result, the adopted relationship is drawn more close to the green line (proposed by Arhin et al. [10]) as the latter is based on data from city streets rather than highways. Varying the PCI in the range of 0 to 100, the resulting IRI in Equation 6.3 falls in the range of 1 to 8 m/km. This IRI range is

sensible according to several references [115, 144].

$$IRI = -\frac{1}{14} \times (PCI - 114) \tag{6.3}$$

The resulting IRI can be used to predict fuel consumption and emissions according to various studies. Generally, one unit (1 m/km) of increase in IRI is reported to bring about 1-3% of increase in fuel consumption [138, 193, 241, 251]. As the value adopted to determine the impact of IRI on fuel consumption and CO_2 emissions will probably have large impacts on the simulation results, two values are adopted in this chapter as a sensitivity analysis, namely the 1% (lower value in the literature) and 3% (higher value in the literature) case.

Apart from the roughness (indicated by IRI), some research also recognises the macroscopic texture (indicated by the MPD) to influence vehicle fuel consumption, as shown in Equation 6.4. However, the contribution of MPD is ignored in this study as (1) the magnitude of MPD is usually lower than the IRI; (2) the coefficient a_1 applied to MPD in a emission model is a tenth of the coefficient a_2 for IRI [241].

$$\begin{aligned} &Tailpipe \quad CO_2 \quad (tonne \quad per \quad 1000 \quad miles) \\ &= a_0 + a_1 \times MPD(mm) + a_2 \times IRI(m/km) \end{aligned} \tag{6.4}$$

Step 3. Calculating link-level emissions based on simulated traffic flow

The unit CO_2 emission per vehicle per road link can be determined in Step 1 and 2. In Step 3, this unit emission is multiplied by the traffic volume (updated at each sub-step) from the traffic mobility simulation module to calculate the total CO_2 emission per link. This step will produce results on the total emissions for each link in the network, which can be used to prioritize the maintenance of certain roads that have the potential to offer the greatest carbon emission savings after receiving maintenance treatments.

Step 4. Calculating network-level emissions

Step 4 is to calculate network-wide infrastructure condition, transport efficiency and CO_2 emission metrics based on the results from previous steps. The metrics include:

1. CO_2 emissions: Annual average daily CO_2 emissions (AAD-CO_2) on the whole or a partition of the network managed by different agencies. The highways are managed by the California Department of Transportation (Caltrans), while the majority of the local roads are managed by the SF Public Works. A small percentage of the local roads are privately managed. According to SF Department of Environment [211], the

average daily GHG emissions (including CO_2, methane and nitrous oxide, etc.) are 4,500 metric tons. This figure includes the estimated contributions of the heavy trucks, which are not considered in this study.

2. Traffic efficiency:

   (a) Annual average daily vehicle hours travelled (AAD-VHT). This indicates how many hours the whole population of travellers spend in the network. Since the travel demand (origin-destination, OD, matrix) is fixed, the smaller the value is, the faster each traveller gets to the destination.

   (b) Annual average daily vehicle kilometres travelled (AAD-VKMT). This metric indicates how long the travellers travel in distance to complete their journeys. A larger value indicates that some travellers are taking longer routes for, e.g., taking the highways outside of the city centre.

3. Infrastructure condition: Average PCI of all streets in the city. This is the usual indicator used by asset managers to evaluate the pavement conditions. It is a numerical scale from 0 to 100, where an average PCI of 0 indicates the all roads in the city are totally deteriorated. An average PCI of 100 indicates that all roads in the city are new. The average PCI of local roads in SF (not including the highways) is usually around 65 to 74 [212].

Interfaces with the traffic mobility and pavement degradation model

The development of the traffic mobility simulation module and the pavement degradation model have been explained in detail in Chapters 3 and 5. This section highlights some modifications that are adopted to facilitate the integration of the two modules for system efficiency and CO_2 emission simulations.

The route choice in the traffic simulation module presented in Chapter 3 is based solely on the link-level travel time at the time of departure. Travellers will choose the routes that minimise their journey times. However, under certain low carbon travel strategies, some travellers will instead travel on the routes that minimise the total journey CO_2 emissions. In such cases, the pavement condition data of each street will be transferred to the corresponding links in the road network graph. The link-level PCI from the pavement degradation module and link-level average speed from the traffic mobility module will be combined to determine the link-level emission factor. This becomes the CO_2 emissions per vehicle traversing a link given the current traffic speed and pavement condition. This link-level CO_2 emission factor

per vehicle will then be used as the link weight and to be used in finding the least emission path using the Dijkstra's algorithm.

For the pavement degradation and maintenance module, it is modified to receive the total link-level traffic emissions. The total link-level emissions are calculated by summing the emissions from all vehicles traversing the link, which in turn depend on the time-varying traffic speed and the current pavement conditions. The total link-level emissions are stored as road attributes in the pavement degradation and maintenance module. Under certain emission mitigation strategies, pavements that can offer the highest reductions in CO_2 emissions have higher priority in receiving maintenance.

Other interactions between the traffic mobility and pavement infrastructure modules are not fully included, notably (1) the delays and congestions due to construction disruptions, (2) the reductions in free flow speed or road capacities on degraded roads and (3) the effect of traffic load on pavement degradation. Construction delays (1) are not considered mainly because of the computational difficulty. Currently, it takes around 6 minutes to carry out traffic simulations for a typical day (around 2 million trips). For an analysis period of 10 years, simulating a different traffic patterns according to the daily construction plan would lead to 300-400 hours of simulation for just one scenario. As a result, the effects of pavement roadworks on the traffic efficiency and network-wide emissions are only carried out in the preparatory analysis for one typical day on three different roads. (2) is not considered as there is no consensus as for the direct impacts on traffic flow speed and capacity from the literature [242]. The exclusion of traffic load effect (3) in the pavement degradation module is based on the conclusion from Chapter 5, which may be a result of the data quality of this study. It is likely that in future studies based on a different dataset, the effect of traffic load will become a more significant explanatory factor than pavement age since last maintenance.

6.3 Preparatory analysis

The simulation of transportation network-level CO_2 emissions relies on a series of assumptions. This section presents some preparatory simulation experiments conducted around two key assumptions. It includes a desktop study that identifies the suitable pavement maintenance method and budget constraints based on reports published by the local agencies. Additionally, as the current simulation speed does not permit a detailed day-to-day analysis of the traffic delays caused by roadworks, a group of smaller simulations are conducted to quantify the effects of such disruptions. Although not conclusive, such simulations can offer some intuition as for how different the results would be if maintenance delays were considered.

6.3.1 Pavement maintenance method and budget

The SF road network is mainly managed by two authorities: the Caltrans which is responsible for the 130 km of highways, as well as the SF Public Works that maintains the 1,400 km of local roads. In the past, there have been numerous studies investigating the environmental impacts of highway segments during the construction, operation and maintenance stages. However, similar works on local roads are rarely seen. This is probably because on local roads, where the traffic volumes are considerably lower than on the highways, the extra roadworks to keep the road smooth are often considered not cost effective if the goal is to reduce the emissions [203]. But in busy urban areas such as SF, over half of the mileages of passenger car journeys are on local roads. The local roads usually have worse conditions than the highways (higher potential of maintenance improvements) and the difficulty in roadworks are also smaller (less traffic disruptions and maintenance costs). Especially considering the carbon emission and mitigation at the network level, the contributions and opportunities from the local roads could become an important component.

As the pavement degradation model is built with data from the local road network, it will continue to be the main focus of the emission reduction analysis. Besides, it also serves to fill in the research gap and improves the understanding of the CO_2 mitigation opportunities from the local roads. For the highways, it is assumed that they enjoy a stable level of maintenance funding and their conditions can be sustained at the PCI of 85. This corresponds to an IRI of 2 according to Equation 6.3 and is in accordance with a Caltrans report stating that the majority of the highways in the Bay Area have IRI between 1.5 and 2.6 [37]. The emissions due to the roughness of the highway part of the network will be calculated based on this PCI value. According to the IRI sensitivity factor identified above, this corresponds to a 1 or 3% increase in fuel consumption and emissions, compared to driving on brand new roads.

For the local pavement network, the initial PCI at the beginning of the analysis (2017) and segment-level degradation rates are borrowed from the model from the Chapter 5. However, the initial average PCI from the model (PCI = 79) does not match the latest observation released by the SF Public Works, which is 74 [212]. Thus, the intercept term are recalibrated to reflect this new information. In the results section, a comparison will be included as for the differences in the outcomes depending on whether the initial PCI of 74 or 79 is used. Also, it is noticed that the PCI degradation rates from Chapter 5 are rather small (on average the PCI reduces by 1 per year). This is probably due to the fact that some small localised pavement repair activities that improved the pavement conditions were not recorded in the degradation dataset, so the pavements appear to degrade slower. To address this possibility, a sensitivity analysis is also included to investigate the differences in outcomes if the actual degradation rates are faster than the regression results from Chapter 5.

For the maintenance of the local roads, several treatment techniques are found from the SF Public Works reports (Table 6.3, [48, 49]). In these documents, street block often serves as the unit of roadworks. A street block is a two-way road segment between consecutive intersections and is usually around 100 metres. The OSM network adopted in this study contains 11,296 blocks of local streets, which is about 12% less than the official figure of 12,855 blocks [49]. This is because the many small streets leading to the entrances of buildings are considered to be service roads and are not included in the traffic simulation network. In this chapter, only the repaving maintenance technique is studied, as the cost of the smaller preservation-type maintenance is cheaper, meaning that they may be carried out more sporadically. In addition, they may not offer significant improvements of pavement roughness as well [238]. Reconstruction works can reset the pavement roughness to the initial value. However, as the agency's budget is limited, the scheduling of the expensive reconstruction works should be a more complex process and should prioritise factors such as the structural safety.

The reports from SF Public Works highlight the importance of budget to pavement maintenance [48, 49]. With a funding level of 218.6 million US dollars for three years from a special bond and other sources, the agency is able to preserve, repave, construct and reconstruct 2,525 street segments in three years. In particular, 148.4 million dollars are from a special bond that will be used specifically to treat 1,389 street segments over three years. In reality, the resurfacing program received a larger share of the fund, allowing 1,423 streets to be resurfaced [49]. As no further breakdown of the funding is available regarding the split between different maintenance techniques, it is estimated based on the unit cost of Table 6.3 for the total number of repaving projects at two funding levels.

In Table 6.3, three types of pavement treatment methods are listed, including the preservation, repaving and reconstruction. The preservation methods are the least disruptive ones, with commonly used techniques being the slurry sealing, crack sealing, etc. These techniques are applied to pavements with medium level of degradation (PCI between 64 and 84) to correct minor defects. The costs associated with pavement preservations are also the lowest. The next type of maintenance treatment is the road repaving, which involves milling off the top layer of the existing road surface and applying a new layer of asphalt. It is applied to pavements that are deteriorated into a PCI of 50 to 63 and the cost of repaving is more expensive than road preservations. When a road is badly deteriorated (PCI $\leq$ 49), reconstruction or similar major interventions are needed. The cost of this type of roadworks is the highest and the traffic disruption time is also the longest. Combining the usual cost estimation in Table 6.3 with the budget information from the local agency [49], the two levels of budget assumed in this study are estimated as the following:

Table 6.3 Pavement treatment categories [48, 49].

Treatment	Applied to PCI	Average cost per block (in US dollars)
Preservation (slurry sealing or crack sealing to extend the life of the pavement)	64 - 84	9,000
Repave (grind off and replace the top two inches of asphalt, sometimes called overlay)	50 - 63	97,800
Reconstruction or resurface with base repair	0 - 49	436,400 (for reconstruction) or 140,000 (for resurface with base repair)

- The "low budget" case: Without sustaining the funding level available with the special bond, the pavement maintenance budget will be $218.6 - 148.4 = 70.2$ million dollars. This is split into three years, equalling 23.4 million dollars per year. As preventative maintenance is relatively cheap and construction/reconstruction is rare, it is assumed that the repaving project takes the majority of budget. Given that it takes 97,800 dollars for repaving a block of street, the low budget case seeks to repair $23,400,000/97,800 \approx 200$ blocks of pavements each year (1.56% of all pavement segments).
- The "high budget" case: If the relatively ample funding income can be sustained as during the special bond funding period in 2011-14, equivalent to $148.4/3 = 49.5$ million dollars of extra funding per year will be available on top of the 23.4 million dollars from the low budget case. This leads to a budget that allows a total of 700 streets to be resurfaced annually, which is identified as the high budget case.

6.3.2 Quantifying impacts to traffic during roadworks

Pavement maintenance could cause traffic disruptions and induce additional CO_2 emissions. As for whether this cause of emissions is significant to gain consideration, different conclusions have been reached in the literature. In Galatioto et al. [83], based on a case study of an inter-urban road in the UK, it is found that the emissions caused by traffic disruptions at roadworks are relatively small (1% of the construction process emissions) unless the

roadworks are carried out during the peak hours or on high traffic volume roads. While studying a 1,600 lane-km sample of asphalt pavements in California, Reger et al. [193] totally ignored the emissions from traffic delays on the ground that highway rehabilitations are usually carried out overnight, so as in Wang et al. [241].

It is desirable to include the roadwork disruption effects into the analysis of this chapter. However, this will make the computation prohibitively expensive giving the current speed of the traffic simulation. Currently, it takes nearly one hour to simulate the hourly traffic distribution of a seven-day week. While in this chapter, it is planned to simulate the traffic over 10 years under many scenarios. As a result, it is impossible to include day-by-day variation of traffic due to the construction work. Given it is important to understand the construction delay and emission increase, several single day simulations are conducted in this section to investigate the roadwork delay effects.

Three roads are chosen from the network, which respectively have the highest, median and lower quantile traffic volumes among all local streets on a typical Friday. The locations of these representative streets are given in Figure 6.4. Simulations are carried out with these streets being closed on Friday (the busiest day of the week) and various network-level metrics are calculated and presented in Table 6.4. The daily vehicle volumes on each street without maintenance disruptions are 302, 5,676 and 94,904, respectively. When maintenance works are carried out on these three representative local streets (in three independent simulations, not simultaneously), the daily traffic volumes on these streets become 0 as expected. For construction works carried out on the streets with low or median levels of traffic, the disruptions to traffic are almost discernible at the network-level compared to the normal operating conditions without maintenance disruptions. If the roadwork is carried out on the street with the highest traffic volume, the disruption effects are more visible, albeit still small. The CO_2 emissions from the local road network or the whole network are around 0.5% higher than the normal operating conditions. The total VHT increases by about 1%, while the changes in VKMT are less obvious.

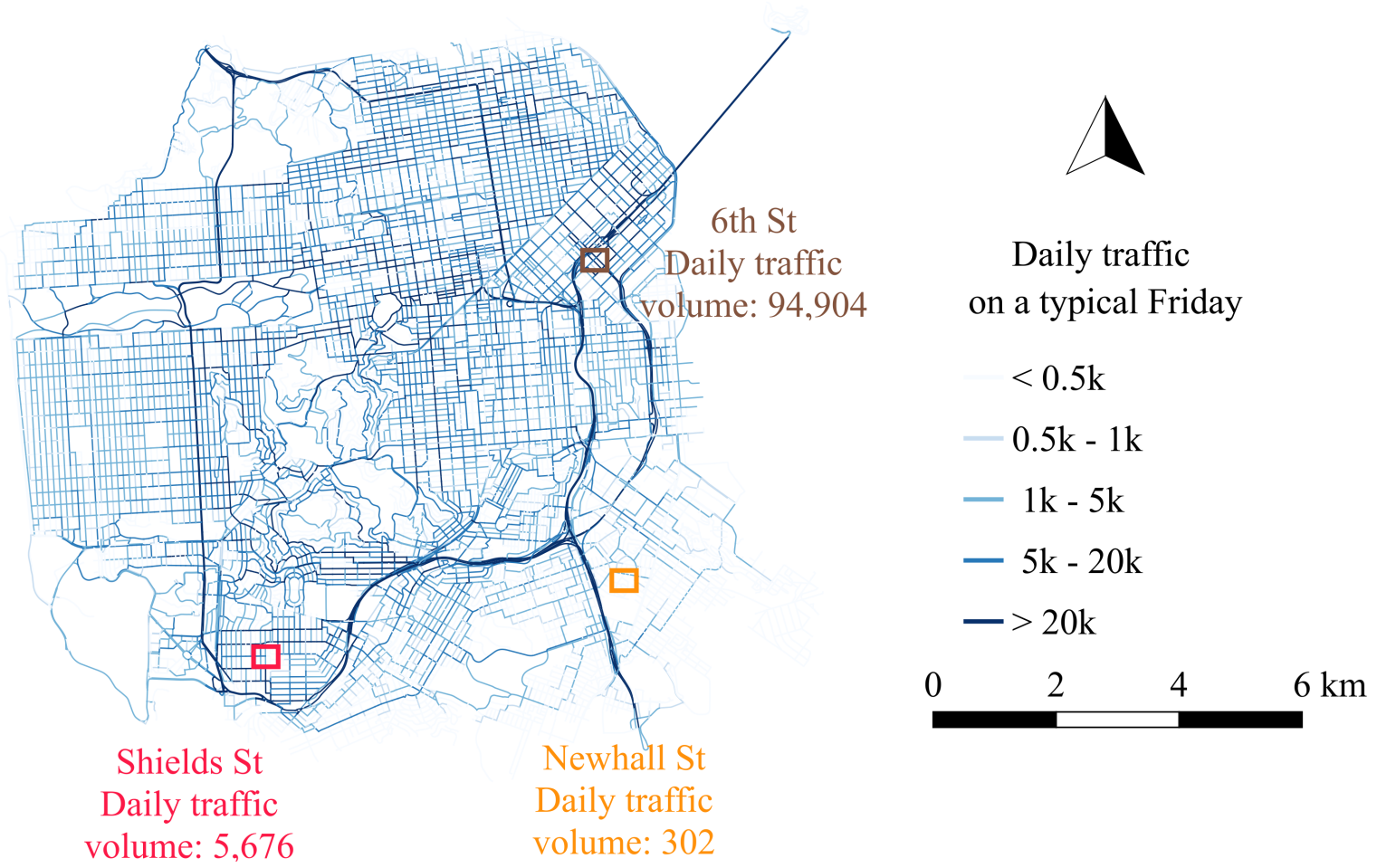


Fig. 6.4 Three representative streets used for construction closure analysis. The numbers are the uninterrupted daily traffic volume (on a typical Friday).

Table 6.4 Road closure disruptions on three representative local streets

Road closure case	No closure	Low traffic road	Mid traffic road	High traffic road
CNN (road ID)		’9580000’	’11808000’	’340000’
Street name	-	Newhall Street	Shields Street	6th Street
District	-	Bayview	Merced Heights	Mission District
Daily traffic volume before closure	-	302	5,676	94,904
Daily traffic volume with closure	-	0	0	0
AAD-CO_2 of local roads (t)	2,506	2,507	2,503	2,520
AAD-CO_2 of whole network (t)	4,710	4,713	4,710	4,732
AAD-VHT of local roads	435,455	435,609	434,973	439,884
AAD-VHT of whole network	795,352	795,266	795,313	802,466
AAD-VKMT of local roads	7,965,807	7,970,672	7,962,343	8,014,892
AAD-VKMT of whole network	18,397,731	18,401,193	18,394,666	18,396,228

It should be noted that in this preparatory analysis, roadwork closures are assumed to last for 24 hours. While not all roadworks are carried out on high traffic volume roads, the long-term effects of roadworks maybe smaller than the last column (closing the busiest road) in Table 6.4. However, on the other hand, maintenance on several sites may be carried out on the same day, compounding the traffic disruptions in the city. As a result, Table 6.4 does not provide a conclusive answer as for the maintenance impacts, including the additional emissions caused by the traffic disruptions. Nevertheless, such results are still useful for understanding the implications of not including the day-to-day maintenance disruptions in the full simulation.

6.4 Simulation set-up

Several scenarios are designed to explore the outcomes in terms of the infrastructure conditions, transport efficiency and CO_2 emissions under different carbon mitigation measures. From the traffic side, a change in driver behaviour is introduced by assuming eco-routing at various participation levels. From the pavement asset management side, an emission mitigation oriented maintenance site selection scheme (eco-maintenance) is compared with the scheme where the maintenance of roads with the worst conditions (PCI-based maintenance) is prioritised. Furthermore, two sensitivity variables are included throughout all simulations, namely the maintenance budget and the influence on vehicle emissions due to the roughness increase (IRI sensitivity). The parameter set-ups of the total 32 simulations are summarised in Table 6.5. Scenarios concerning traffic growth are presented additionally in Section 6.6.

6.4.1 Scenario 1: PCI-based maintenance

This scenario is designed to reflect a maintenance strategy that does not prioritise the reduction of the vehicle CO_2 emissions due to pavement roughness. The actual pavement selection is a complex decision coordinated between the pavement asset managers, utility companies and other agencies. Besides, the distribution of the improvement opportunities need to consider the equity between commercial and residential areas. According to the City and County of San Francisco [48], the street repaving program will prioritise the transit and biking networks, PCI scores, functional classifications, project readiness, equitable distributions and public complaints. Many of these realistic issues cannot be fully reflected in the simulations. So Scenario 1 is designed as a simplistic representation of the status quo, namely the pavements are selected for maintenance solely based on the PCI score. In this scenario, the city-wide average pavement condition is expected to improve given sufficient budget.

Table 6.5 Simulation scenarios

No.	Maintenance priority	Eco-routing percentage	Budget (streets/year)	IRI sensitivity (CO_2 increase per 1 m/km IRI)
1a	Improve PCI	No	200	1%
1b			700	1%
1c			200	3%
1d			700	3%
2a	Reduce CO_2 (eco-maintenance)	No	200	1%
2b			700	1%
2c			200	3%
2d			700	3%
3a	Improve PCI	10%	200	1%
3b			700	1%
3c			200	3%
3d			700	3%
3e		50%	200	1%
3f			700	1%
3g			200	3%
3h			700	3%
3i		100%	200	1%
3j			700	1%
3k			200	3%
3l			700	3%
4a	Reduce CO_2 (eco-maintenance)	10%	200	1%
4b			700	1%
4c			200	3%
4d			700	3%
4e		50%	200	1%
4f			700	1%
4g			200	3%
4h			700	3%
4i		100%	200	1%
4j			700	1%
4k			200	3%
4l			700	3%

As for the traffic behaviours in this scenario, it is assumed that travellers still choose the fastest route. In other words, there is no change in travellers' behaviours compared to Chapters 3 and 4 of the thesis.

6.4.2 Scenario 2: CO_2 emission mitigation oriented maintenance site selection (eco-maintenance)

The second scenario is designed to quantify the CO_2 emission reduction benefits of a new pavement management scheme. Specifically, the maintenance site selection is assumed to focus on repairing the pavements for the purpose of reducing the use phase vehicle emissions caused by pavement degradation. The limited resources are allocated to pavement segments that can offer the greatest CO_2 savings. Pavement segments receiving treatments usually have low PCI (as in Scenario 1), as well as high traffic volumes. Maintaining pavement segments that satisfy both conditions (low PCI, high volume) leads to the highest reductions in terms of CO_2 emissions. In Scenario 1, the roads with the most severe deteriorations are selected for maintenance. However, if only low numbers of vehicles use these streets, the potential of CO_2 reduction will not be significant. Thus in Scenario 2, some less used poorly deteriorated streets may not be treated with high priority.

In this scenario, travellers are assumed to keep the route choices that minimise the journey times. As a result, the eco-maintenance scenario is not expected to alter the traffic behaviours of the travellers as well.

Scenarios 1 and 2 focus on the effects of different pavement maintenance strategies. Such analysis relies heavily on the assumptions of the underlying pavement degradation process. In Chapter 5, a spatially correlated pavement degradation model has been proposed. But Chapter 5 also shows that, in reality, the pavement condition data are highly scattered and the degradation trends hard to be fully captured by any model. To account for the possible deficiencies in the pavement degradation model, a parameter sensitivity analysis will be carried out. This includes an investigation of the following pavement degradation model parameters:

- Starting conditions: As stated at the beginning of this chapter, the initial condition obtained from the degradation model for the first year of analysis (2017) is 79. However, the record from the SF Public Works shows a PCI of 74 on the same year [212]. Recalibrating the model to match this new information, the starting conditions of all pavement segments are set to 5 less in this chapter. It will be shown how this shift in the starting conditions affects the simulation results.

- Degradation rates: Based on the pavement degradation model obtained in Chapter 5, on average the pavement PCI only drops by about 1 point per year. This degradation rate is considerably lower than the general perception. Generally, given a pavement design life of 20 years, degradation from PCI of 100 to around 40 or 50 is expected to happen. There are multiple reasons for the low degradation rates to be obtained in Chapter 5. First of all, in the pavement condition dataset used for the regression analysis, small maintenance repairs that improve PCI slightly are not documented, making the degradation rates appear to be slower. On the other hand, as the pavement PCI tends to drop significantly in a short time period after construction (usually within the first few months), the intercept term (age zero condition) in the regression model is not fixed to 100. Without this constraint, the slope of the regression model also becomes smaller. To test if a faster degradation rate makes a difference on the outcomes of eco-maintenance, a sensitivity comparison on the degradation rate is also included.

- Maintenance gain/quality: It is not guaranteed that maintenance treatment can always bring the pavement PCI back to 100. Research has suggested a partial recovery of PCI due to maintenance such as pavement overlay [65]. As a result, the potential of PCI improvements through maintenance is also investigated as for its effect on the simulation results.

6.4.3 Scenario 3: eco-routing

In Scenario 3, some travellers start to incorporate the low carbon criteria into their route choices. In the traffic simulation developed in Chapter 3, all travellers are assumed to choose to travel on the fastest routes. While in this chapter, it is assumed that some travellers become more concerned about the footprint of their journeys and choose the routes that will generate the least CO_2 emissions (eco-routing). In this scenario, link attributes, including the speeds (congestion levels), the gradients as well as the pavement conditions will all affect the traffic distributions to various degrees. In particular, the congestions and pavement conditions are time-varying and require the coupled traffic and pavement degradation simulations to be carried out sequentially for the duration of the study period.

Three eco-routing sub-scenarios are evaluated, with 10%, 50% and 100% out of all the trips in each hourly travel demand being assumed to switch to the eco-friendly path. The routing criteria (shortest time or least emission) are randomly assigned to drivers at the beginning of the simulation, but the total percentages of eco-routing trips (10%, 50% and 100%) are consistent with the sub-scenario specification. To investigate the effect of the random splitting of eco and non eco-routing trips, repeated simulations with various random

seeds are included. Compared with the eco-maintenance strategy, eco-routing represents more of a bottom up behaviour change voluntarily adopted by individual drivers, but this behaviour can also be incentivised through policies such as emission charges. The current trip-based nature of the traffic simulation is not sophisticated enough to incorporate complex traveller behaviours, such as stratifying the population and linking the route choices with the socio-demographic status. This extra layer of complexity can be explored through activity-based traffic models in future research.

6.4.4 Scenario 4: eco-maintenance and eco-routing combined

Scenario 4 investigates the outcomes based on the combined eco-maintenance and eco-routing strategies that were introduced separately in Scenarios 2 and 3. This scenario is designed to quantify the combined values of implementing different emission mitigation approaches from two distinct perspectives (traffic operation and pavement asset management). In particular, due to the existence of interactions between the traffic operation and pavement asset management, this scenario is important for evaluating the outcomes of simultaneously implementing multiple carbon mitigation measures in the same system. The following interactions between the traffic operation and pavement maintenance are considered:

1. With eco-maintenance, the conditions of the local roads will be different from those obtained with the PCI-based maintenance scheme (in Scenarios 1 and 3). This change in the pavement conditions will lead to an alteration in the link-level emission factors. As some road links become more eco-friendly while others less, this might attract or dissuade the eco-routing travellers from using them.

2. With eco-routing, the traffic distributions will also be different from those obtained from time-based fastest routing. Although from Chapter 5 it has been decided that these variations in the traffic distributions will not change the pavement degradation trend (pavement degradation model uses age as the explanatory variable), they will affect the maintenance site selections under the eco-maintenance scheme. If eco-routing leads to more frequent use of certain roads, they will also receive higher priority in the eco-maintenance process.

To quantify the strength of such interactions, the impacts on CO_2 emissions, VHT, etc. of the combined measures will be calculated. It will be compared to see if they are higher, lower or equal to the simple sums of the outcomes when different measures are implemented independently.

6.4.5 Other sensitivity factors

The calculation of CO_2 emissions from the transport system relies on many assumptions as introduced in Section 6.2. Among these assumptions, some are of higher certainty than others, such as the widely cited speed-emission relationship and the PCI-IRI relationship. However, the validity of some other parameters are hard to determine. Specifically, there is no strong consensus as for how much extra CO_2 emissions are caused by a unit increase in pavement roughness (IRI sensitivity), yet this factor has a foreseeable significant influence on the analysis outcomes. As a result, two levels of IRI sensitivity are included in the simulations as a sensitivity parameter: when the pavement IRI increases by 1 m/km, the CO_2 emissions are assumed to grow by 1% as the lower bound or 3% as the higher bound.

The budget level is another sensitivity factor that is carried out throughout all scenarios. The low budget case allows the repairing of 200 street blocks per year, while in high budget case, 700 street blocks can be repaired annually. Given the total number of 11,296 local street blocks, the pavement segments receive maintenance treatments in every 56 years on average in the low budget case, almost twice the pavement service life of 20 to 30 years. While in the high budget case, the pavements are maintained in every 16 years on average.

The traffic growth rate is not considered as a sensitivity factor until in Section 6.6. The reason for not including the increasing demand in the majority parts of the simulations is to isolate the effects of traffic growth on the CO_2 emissions and traffic efficiency measures from the effects of the maintenance and routing strategies. In Section 6.6, the analysis is extended to include the traffic growth rate according to the projections by the local transport agencies [54].

6.4.6 Other simulation set-ups

Due to the computational efforts involved, the study period of each scenario is assumed to lasts for 10 years in most cases. In the sensitivity analysis related to the pavement degradation model parameters, the study period lasts for 20 years to demonstrate the longer-term effects. Throughout the analysis (in this section), the travel demand is assumed to be fixed, i.e., no annual traffic growth. This assumption of fixed travel demand is made so as to separate the environmental effects of pavement management and traffic operation from those caused by the increasing levels of traffic. In addition, forecasting the growth of different types of vehicular traffic (petrol cars, diesel cars, alternative fuel vehicles and public transits) and converting the growth to trips (attaching origins, destinations and departure times) are significantly more complex. However, in order to understand the impact of potentially higher

travel demand, the traffic growth scenarios are included in Section 6.6 as stated in the above section.

All simulations are conducted on the HPC facilities. It takes around 6 minutes to simulate the traffic of a typical day, but the actual queuing time can be much longer. Simulations of different years that belong to the same scenario (same row in Table 6.5) need to run sequentially, as the route choices and maintenance site selections of Year 2 depend on the traffic distributions and pavement conditions of Year 1, and so forth. To reduce the running time (as well as the queueing time on the HPC), only the traffic on a representative day is simulated for each year in the study period. Among the seven days of a week, Wednesday is chosen as the representative day as its total number of trips (1.6 million) is the closest to the daily average of a typical week. It should be noted that in Chapter 3, traffic simulations are conducted for a whole week. While in Chapter 4, only the Friday traffic is simulated to reduce the computational burden while investigating the impacts of the probe penetration rates and probe data variability. The reason to choose Friday for the analysis in Chapter 4 is because the travel demand on Friday is the highest, so clearest distinctions can be expected regarding the outcomes in traffic efficiency due to the different levels of probe data availability and variability. In this chapter, the Friday travel demand is used in Section 6.3.2, for the same reason as it is being used in Chapter 4, namely to estimate largest possible impacts brought by roadworks. However, for long-term simulations as those in this section, it is important to make sure that the traffic in the selected day is representative of the average case. It is because of this requirement that Wednesday is chosen as the representative day in the long-term simulations.

Simulations of the maintenance works are carried out after the traffic simulations. In the maintenance simulations, the roads are ranked according to their priorities in receiving treatments. The top ranked pavements within the budget allowance are selected for maintenance and their conditions are subsequently improved. Emissions due to the traffic interruptions during the roadwork phase are ignored due to the simulation time constraint. The impacts of construction closures can be referred to in Section 6.3.2 through the three road closure examples.

6.5 Results

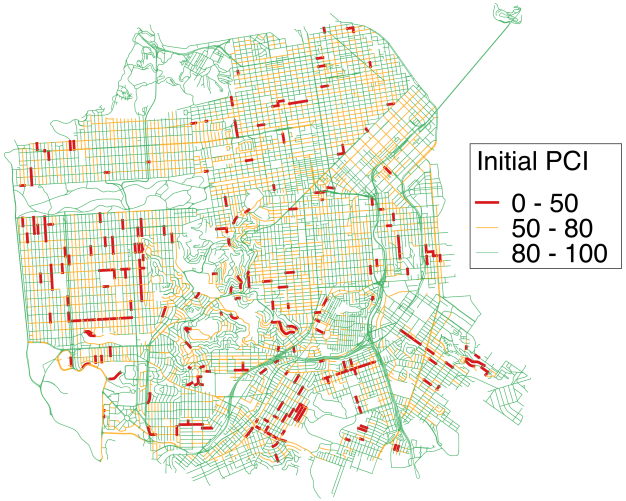
Results of the integrated traffic and pavement degradation/maintenance simulations are presented in this section. The results are interpreted individually as for the effects of the eco-maintenance strategy alone (Section 6.5.1), the eco-routing strategy alone (Section 6.5.2) and then the combined eco-maintenance and eco-routing strategies (Section 6.5.3).

6.5.1 Eco-maintenance

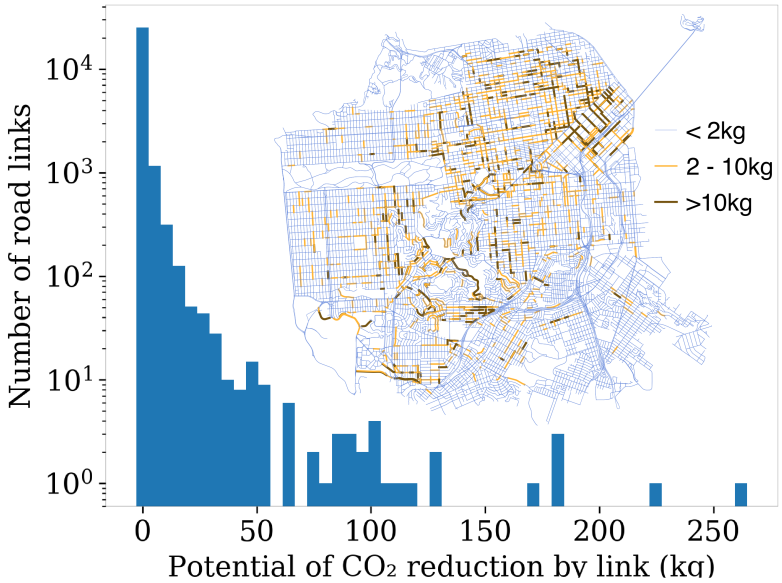
The features of eco-maintenance

In the eco-maintenance scenario, the road pavements with greater potential in CO_2 emission reduction have higher priority to receive maintenance treatments. This means that the maintenance decisions are no longer only based on the pavement conditions, but on a combination of PCI and the traffic volumes that the streets carry. In Figure 6.5 (a), it is shown that the pavement condition at the beginning of the analysis period, where red lines denote pavement segments with low PCIs. Under the PCI-based maintenance site selection scenarios, these red streets receive treatments with higher priority. Figure 6.5 (b) plots the histogram and the spatial distribution of the CO_2 emission reduction potential of all the streets at the beginning of the analysis. It can be seen that the majority of the streets do not offer a CO_2 reduction larger than 5 kg (annual average daily value) if maintenance works are carried out on them. In the eco-maintenance scenario, the maintenance of the black coloured streets in Figure 6.5(b) is prioritised over the others. These streets are not the same as those deemed critical in the PCI-based maintenance scenario as shown in Figure 6.5(a).

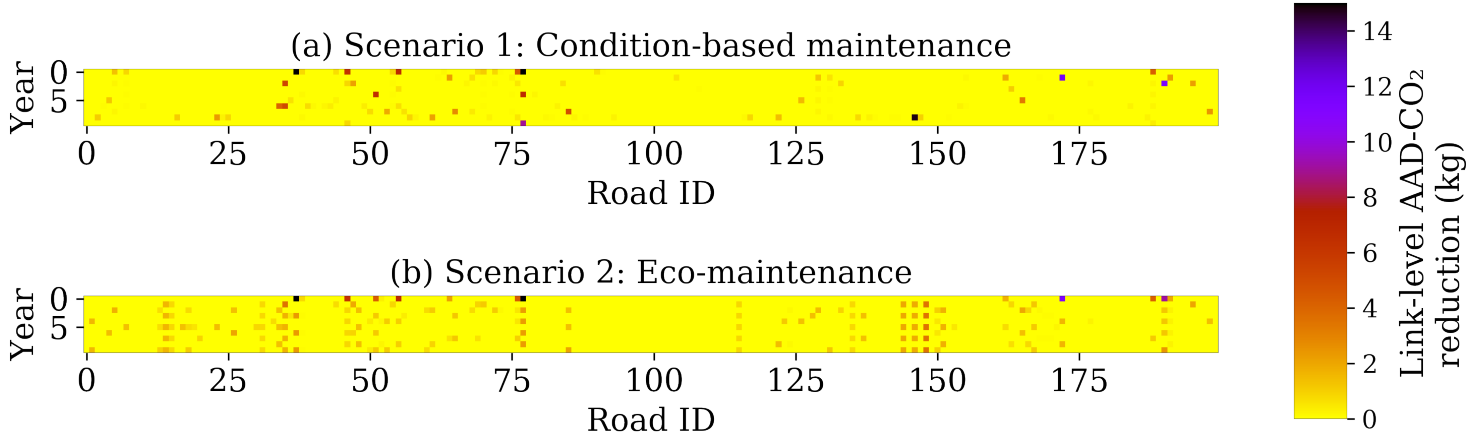
Figure 6.5(c) compares the features of the PCI-based maintenance site selections and the emission reduction-based eco-maintenance more directly in a space time plot. For each of the two matrix plots, the horizontal direction denotes the road ID and the vertical direction shows progress of the simulation years. Each cell, coloured from yellow to dark purple, indicates the actual reduced CO_2 emissions after each maintenance cycle. As long as the colour of a cell is not yellow, it indicates that the corresponding street has been maintained in the particular year. Only results of the first 200 streets are shown here for clarity. For condition-based maintenance, it can be seen that the maintenance pattern is rather sporadic and almost no streets are repaired more than once in the ten year period. While contrastingly, in the eco-maintenance case, many streets are revisited every other year. These streets usually have faster degradation rates and relatively high traffic volumes, so in the eco-maintenance simulation it is necessary to keep their conditions as good as possible to achieve the maximum reductions in CO_2.



(a) Initial Pavement Condition Index (PCI).



(b) Initial CO_2 reduction potential per day by road link.



(c) Maintenance Patterns.

Fig. 6.5 Locations and patterns of condition-based and eco-maintenance.

Impacts of eco-maintenance on CO_2 emissions and other network-level performance metrics

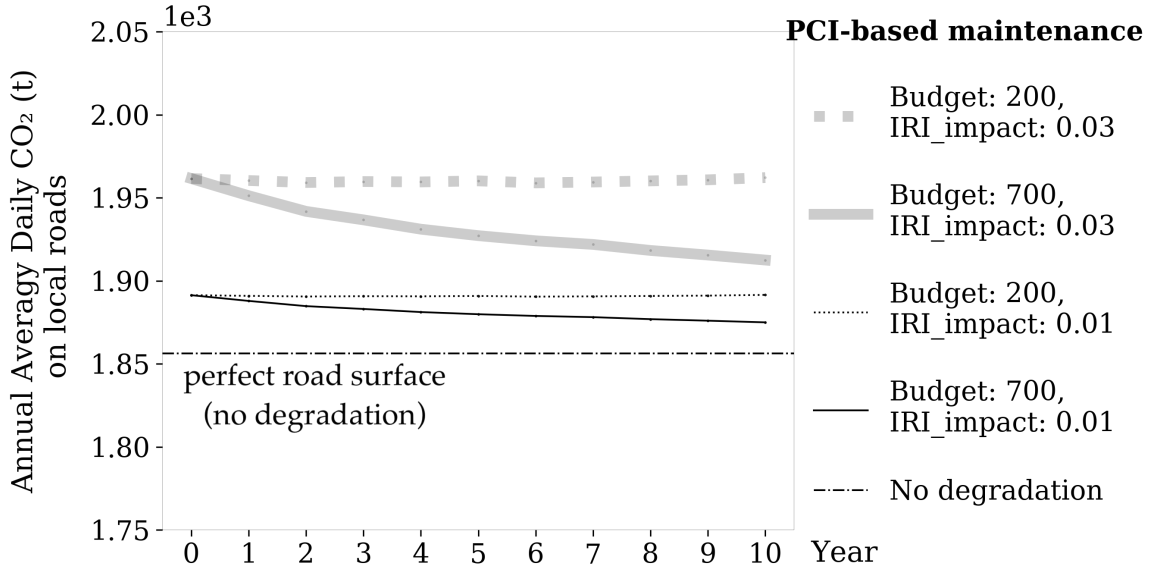
As it is assumed that the travel demand is fixed, there is no change in travellers' route choices and traffic distributions over the study period when only PCI-based maintenance (Scenario 1) and eco-maintenance (Scenario 2) are implemented (without eco-routing). Also, as it is assumed in the simulation set-up that the highways have a fixed roughness over time, the emissions from the highway part of the road network are also constant in these two scenarios. Emissions from the local network over the 10 year study period are shown in Figure 6.6(a) and (b). Figure 6.6(a) displays the AAD-CO_2 of the local roads while implementing the PCI-based maintenance. Figure 6.6(b) shows the results under the eco-maintenance scenario. Different line styles represent the results based on each sensitivity factor combinations (thick lines: IRI sensitivity of 3%; thin lines: IRI sensitivity of 1%; solid lines: high budget level; thin lines: low budget level). Furthermore, to assist the interpretation of the results, the CO_2 emission level assuming no road degradation is plotted as the horizontal black dash-dot lines. This "perfect road emission" is the baseline to quantify the extra emissions caused by the pavement degradation. Given that the traffic distributions do not change in these two scenarios, the "perfect road emission" lines are constant throughout the analysis period.

From the two plots, it can be seen that, under the PCI-based maintenance scenario (Figure 6.6(a)), the budget level plays an important role in controlling the CO_2 emissions from the local road network. Given a low annual budget that allows only 200 street blocks to be maintained each year, the emissions cannot be reduced over time. If the budget increases so as to allow 700 blocks of roads to be maintained annually, the CO_2 emissions can reduce gradually each year. For the eco-maintenance scenarios (Figure 6.6(b)), the emissions from the local road network reduce under both budget levels. In addition, the downward trend is more rapid in the first few years, i.e., approaching the "perfect road emission" line faster than the PCI-based maintenance approach.

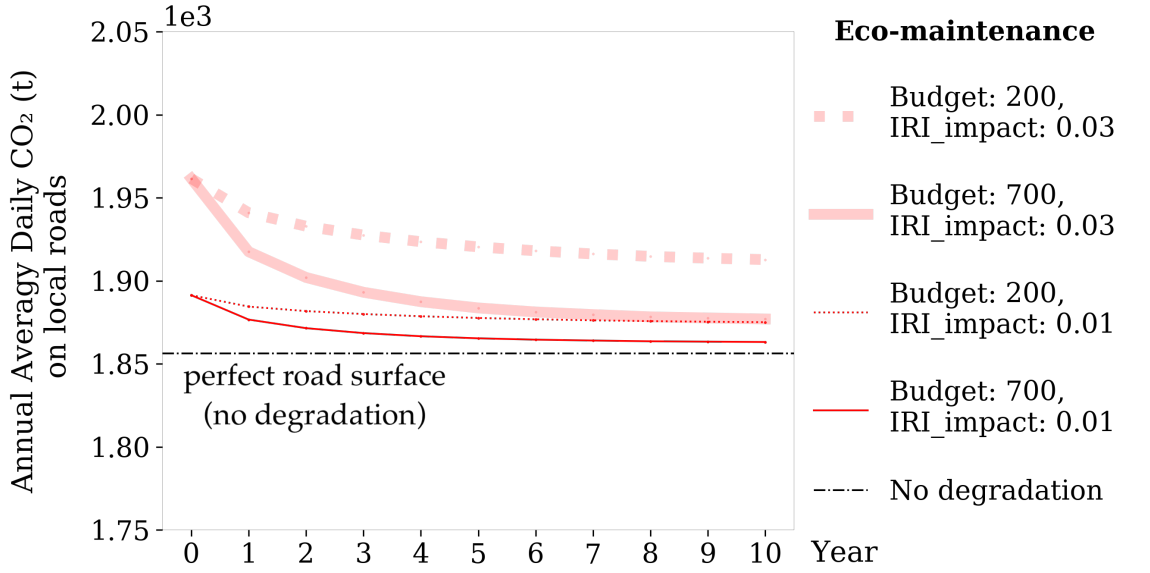
The IRI sensitivity factor also affects the magnitude of the CO_2 emission savings. When the IRI sensitivity factor is low, namely the influence of pavement roughness on the vehicle fuel economy is low, all the numbers will scale down. It can be seen in Figure 6.6 that when the IRI sensitivity factor is 0.01 (1 m/km increase in IRI causes 1% of additional fuel consumption and CO_2 emissions), the local road network emissions are much closer to the "perfect road emission" line. As a result, any difference in budget or maintenance site selection strategies is also less obvious.

Quantitative results of the PCI-based maintenance and eco-maintenance scenarios are given in Table 6.6. Given the IRI sensitivity factor of 3%, increasing the maintenance budget from 200 to 700 street blocks per year will lead to 50 t (2.5%) reduction in AAD-CO_2 from

the local roads at the tenth year under the PCI-based maintenance scenario. The same amount of savings could be achieved by switching to eco-maintenance (with low budget level). In addition, given the IRI sensitivity factor of 3%, 1.9% cumulative reduction in local roads CO_2 over the 10 year study period is obtained after implementing eco-maintenance compared to the PCI-based maintenance. This CO_2 reduction is only 0.6% if the IRI sensitivity factor is 1%, i.e., when pavement roughness does not have a significant influence on the vehicle fuel economy. As no change in traffic demand and routing behaviour is assumed, the network efficiency metrics, including the AAD-VHT and the AAD-VKMT, are not sensitive to the maintenance strategies, budgets, IRI sensitivity factors, etc. Thus all the AAD-VHT and AAD-VKMT values are constant in Table 6.6. However, the average PCI of the local roads are influenced. It is found that the average PCI of the local roads in Scenario 1 (PCI-based maintenance) is 3 to 7 points higher than in Scenario 2. The low budget case is not able to sustain the pavement conditions in both scenarios, while in the high budget case, the pavement conditions improve faster under the PCI-based maintenance site selection.



(a) Scenario 1. Without eco-maintenance.



(b) Scenario 2. Eco-maintenance prioritizing reducing CO_2.

Fig. 6.6 AAD-CO_2 on local road network under Scenarios 1 (prioritise PCI improvement) and 2 (prioritise emission reduction).

Table 6.6 Network-wide metrics in Scenario 1

Year	Year 1				Year 10			
Budget	200		700		200		700	
IRI sensitivity	0.01	0.03	0.01	0.03	0.01	0.03	0.01	0.03
Local AAD-CO_2 (t)	1,890	1,960	1,887	1,951	1,891	1,962	1,874	1,912
Total AAD-CO_2 (t)	3,818	3,931	3,815	3,922	3,821	3,939	3,804	3,889
Local AAD-VHT [1]	337,919	337,919	337,919	337,919	337,919	337,919	337,919	337,919
Total AAD-VHT [2]	649,074	649,074	649,074	649,074	649,074	649,074	649,074	649,074
Local VKMT [3]	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245
Total VKMT [4]	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123
Local roads PCI	74	74	77	77	73	73	86	86

Note: [1-4] Given this scenario does not consider the changes in traffic demand and routing behaviour, the traffic efficiency metrics, including the AAD-VHT and AAD-VKMT, do not change in the simulated years.

Table 6.7 Network-wide metrics in Scenario 2

Year	Year 1				Year 10			
Budget	200		700		200		700	
IRI sensitivity	0.01	0.03	0.01	0.03	0.01	0.03	0.01	0.03
Local AAD-CO_2 (t)	1,884	1,940	1,876	1,917	1,875	1,912	1,863	1,876
Total AAD-CO_2 (t)	3,812	3,912	3,804	3,889	3,804	3,889	3,792	3,854
Local AAD-VHT [1]	337,919	337,919	337,919	337,919	337,919	337,919	337,919	337,919
Total AAD-VHT [2]	649,074	649,074	649,074	649,074	649,074	649,074	649,074	649,074
Local AAD-VKMT [3]	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245	6,359,245
Total AAD-VKMT [4]	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123	15,869,123
Local roads PCI	74	74	75	75	70	70	79	79

Note: [1-4] Given this scenario does not consider the changes in traffic demand and routing behaviour, the traffic efficiency metrics, including the AAD-VHT and AAD-VKMT, do not change in the simulated years.

Sensitivity to pavement degradation model parameters

As the eco-maintenance analysis is probably sensitive to the adopted degradation model parameters, three sensitivity analyses are carried out to understand the effects of varying these parameters, including the initial pavement conditions, the maintenance "quality" and the pavement degradation rates. It should be noted that these sensitivity parameters are specifically related to the pavement degradation model and should not be confused with the global sensitivity parameters for all the scenarios (the budget level and the IRI sensitivity).

First of all, Figure 6.7 shows the impact of offsetting the initial average PCI from 79 to 74. This offset is made so as to match the latest PCI observation in 2017 by the SF Public Works. From Figure 6.7 it can be seen that, while exerting this offset, the estimated AAD-CO_2 from the local road networks are about 20 t, or 1%, higher in the first year. But as the roads get maintained over the years, this difference becomes smaller. Overall, this offset (or the initial pavement conditions used for analysis) indeed has an impact on the evaluation of eco-maintenance. When the average road PCI at the start of the analysis is 74, the cumulative savings in CO_2 emissions on local roads due to eco-maintenance are estimated to be 1.9%. But when the starting PCI is higher (at 79), the cumulative savings of eco-maintenance reduce to 1.5%. This result is in agreement with the intuition that the eco-maintenance strategy becomes less effective when the pavements are in better conditions.

Secondly, it is assumed in the simulations that pavement maintenance (repaving) can bring the pavement conditions back to the perfect state. However, depending on the construction quality and method, the PCI of newly repaved roads may be lower than 100 [65]. Figure 6.8 shows the CO_2 emissions from the local network by varying the maintenance quality or standard. The purple lines are the results when the pavement maintenance recovers 100% of the loss of PCI due to degradation. Under this assumption, the CO_2 emissions from the local network are indeed the lowest. The green lines assume that pavement maintenance can recover 75% of the PCI loss due to degradation. For example, given a pavement segment with a PCI of 50 before repair, maintenance carried out on this segment will bring the PCI back to $(100-50)\times 75\% + 50 = 87.5$. While if the condition before maintenance is 20, the pavement PCI after maintenance becomes $(100-20)\times 75\% + 20 = 80$. In this case, the maintenance is not as effective as the purple lines, when the PCI can always reach 100 after maintenance. As a result, compared with the purple lines, the CO_2 emissions on the local road network are higher for the green lines, when maintenance only recovers 75% of the lost PCI.

The yellow lines in Figure 6.8 show the results assuming that maintenance works can only recover half of the loss of PCI due to degradation. Under this assumption, despite the relatively high funding level (700 blocks per year), the CO_2 emissions are not able to

be reduced through the PCI-based maintenance. While if eco-maintenance is adopted, the AAD-CO_2 emissions will gradually reduce but are still 1-1.5% higher compared to cases when the maintenance works are more effective. This result indicates that the benefits of eco-maintenance is dependent upon the "quality" of the roadworks.

The third sensitivity parameter related to the pavement degradation model is the degradation rate. As minor maintenance activities are not recorded in the the SF pavement PCI dataset, the inferred street-specific PCI degradation rates may be slower than reality. To account for this, three levels of degradation rates are studied, where the pavement degradation rates are assumed to be the same, 3 or 5 times faster than the regression results in Chapter 5. In Figure 6.9, the purple lines show the CO_2 on the local road network with the original degradation rates given by the spatial pavement degradation model in Chapter 5 (the same results as in Figure 6.6 with a maintenance budget of 700 blocks per year and IRI sensitivity factor of 3%). The green and yellow lines show the results when the original pavement degradation rates are multiplied by a factor of 3 or 5, but under the same settings otherwise. When the degradation rates are faster, the CO_2 emissions are also higher. In fact, when the pavement degradation rates are 3 times faster than the original value (PCI reduces by about 3 points per year on average), the PCI-based maintenance (green solid line) almost has no effect on CO_2 emission reduction. While the eco-maintenance site selection can lead to lower CO_2 emissions over the study period, but overall the AAD-CO_2 are still 2% higher than the original case. With a pavement degradation rates 5 times faster than the original value, i.e., PCI loss of around 5 points per year on average, neither PCI-based or eco-maintenance site selection is effective in controlling the overall emissions. This is because the annual loss of PCI due to natural degradation exceeds the speed of repair and a higher budget is required to curtail the growth of pavement roughness induced emissions from the local road network. Also from this set of results, the benefits of eco-maintenance are more obvious when pavement degrades faster: the 10 year cumulative savings in CO_2 emissions from the local roads after implementing eco-maintenance are 1.9%, 2.4% and 2.9% compared to PCI-based maintenance, when the pavement degradation rates are the same, 3 or 5 times faster than the regression results based on the SF pavement condition data.

6.5.2 Eco-routing

Under the eco-routing scenarios, certain percentages (10%, 50% or 100%) of travellers are randomly selected to use the routes that have the minimum journey CO_2, while the rest of the travellers continue to use the shortest time routes. The link-level travel time and carbon emissions are updated after each traffic simulation sub-step based on the traffic volume and speed from this sub-step. To follow the least emission routes, journeys priorly taken on the

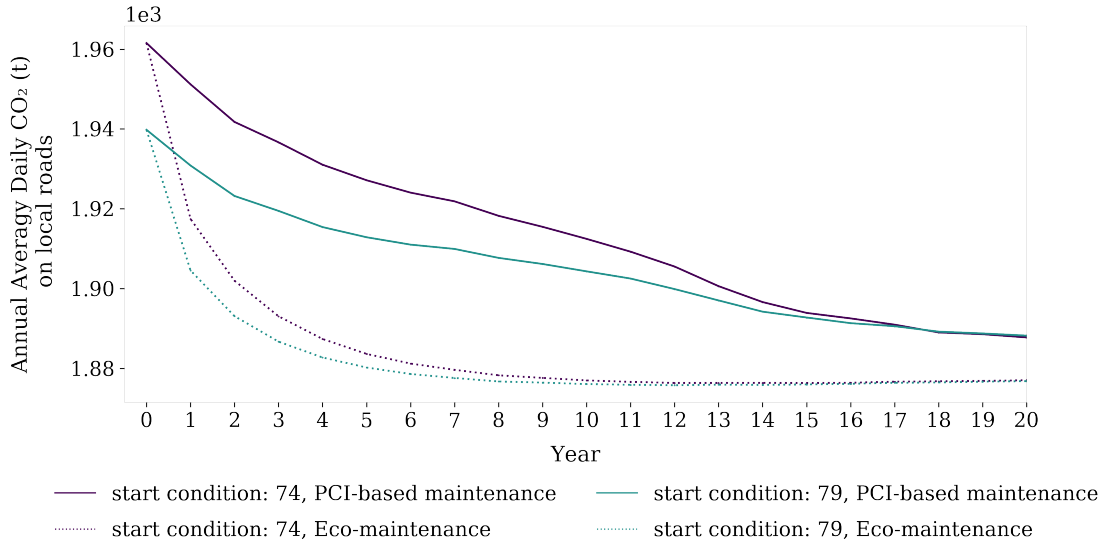


Fig. 6.7 AAD-CO_2 on local road network with or without re-calibration of the initial pavement condition.

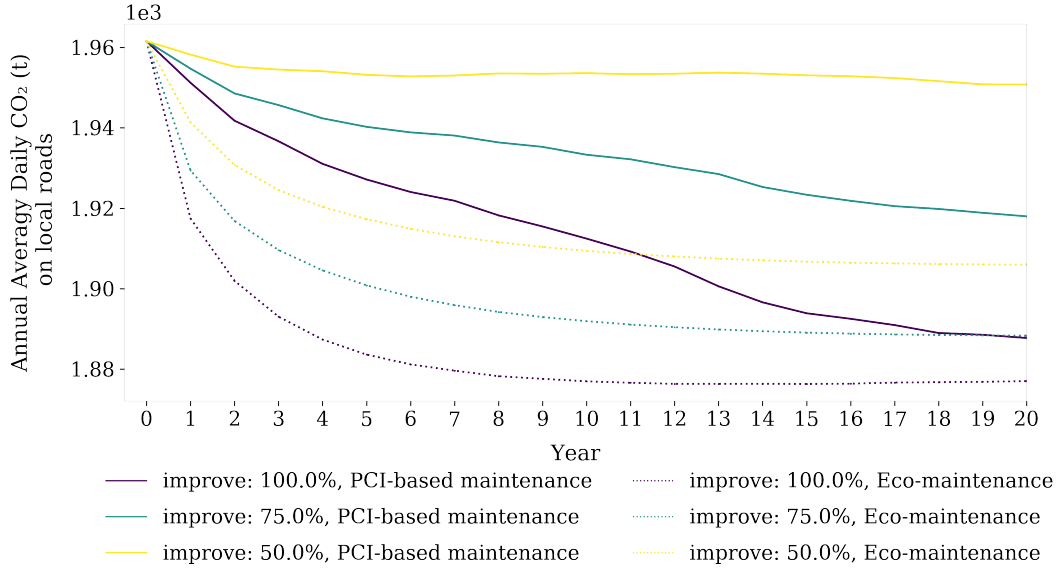


Fig. 6.8 AAD-CO_2 on local road network with different maintenance gains.

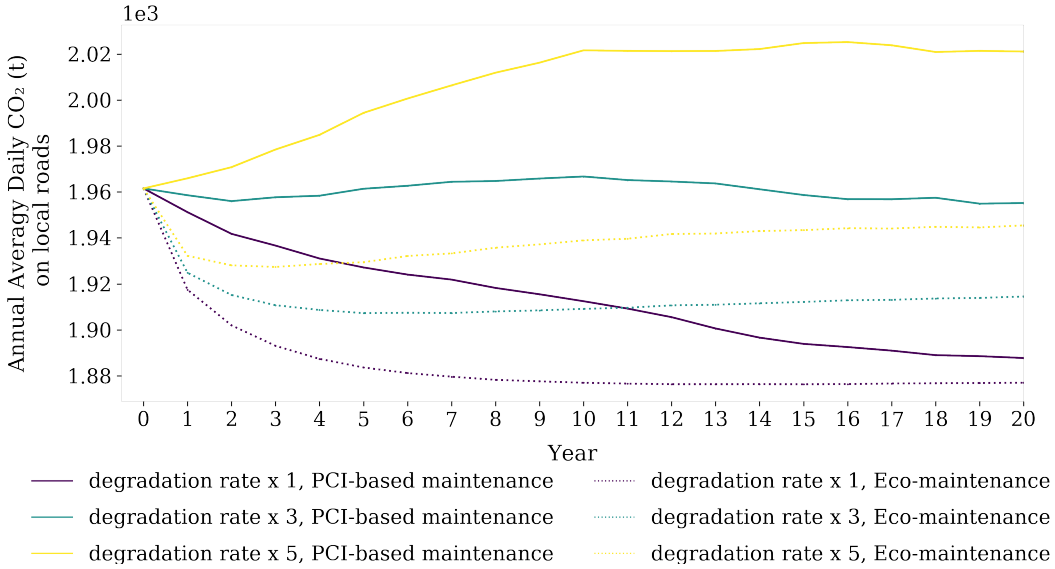


Fig. 6.9 AAD-CO_2 on local road network with different pavement degradation rates.

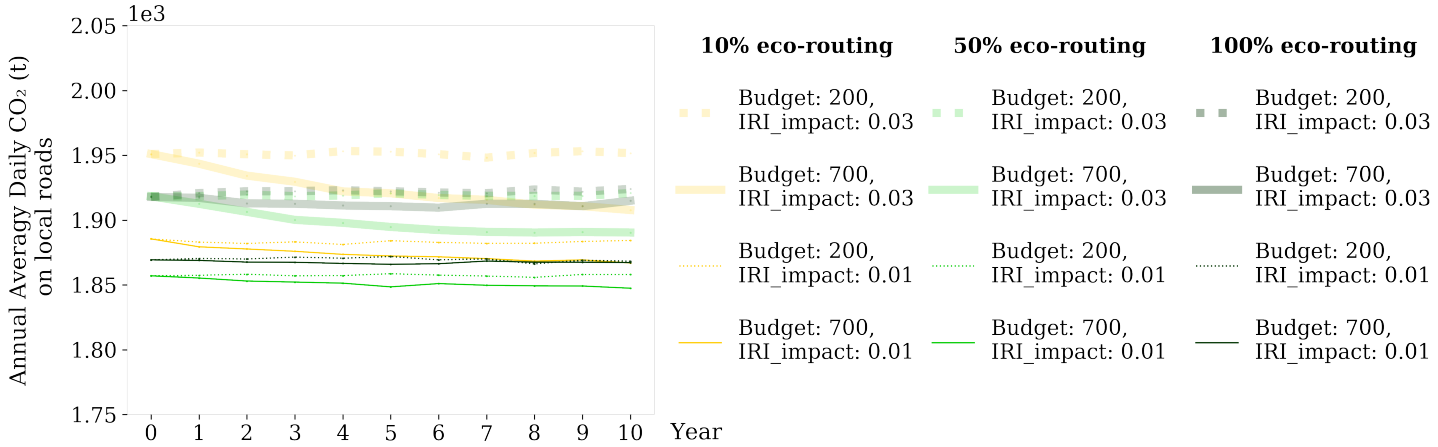
highway may move to the local network and vice versa, so the emissions from the highways are not constants as in Scenarios 1 and 2.

Impacts of eco-routing on CO_2 emissions and other network-level performance metrics

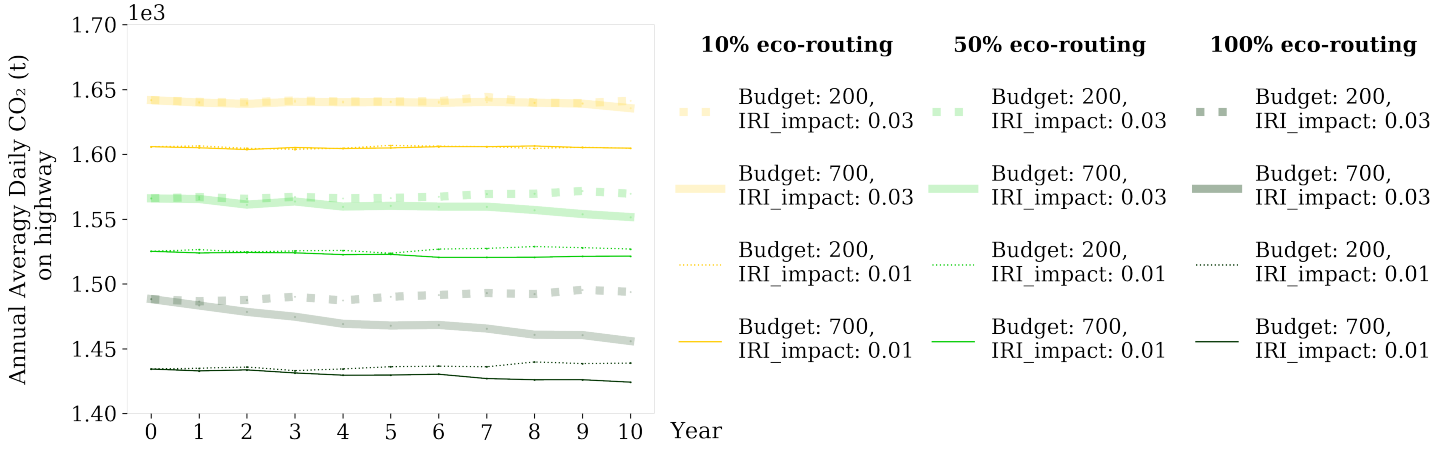
Figure 6.10 shows the emissions from the local road network as well as the highways separately under the eco-routing scenarios. Figure 6.10 (a) and (b) are plotted with different absolute values for the vertical axes but nevertheless the ranges and scales of the two y-axes are the same.

As the percentages of eco-routing travellers increase, the emissions from the highway part of the network reduce significantly. For example, given the IRI sensitivity factor of 3% and a low maintenance budget, the AAD-CO_2 from the highways reduce by 4.5% or 9.4% in the first year when the participation rates of eco-routing trips increase from 10% to 50% or 100%. The underlying reason of this reduced emissions from the highways is that travellers are choosing shorter routes by giving up the longer but faster detours through the highways. With an IRI sensitivity factor of 3% and a low budget level, the AAD-VKMT on the highways reduce by 4% or 11% in the first year when the eco-routing ratio increases from 10% to 50% or 100%. At the same time, the AAD-VKMT on the local network increases slightly by 1% or 2% due to the combined effects of shorter trip distances and increased use of local roads. The AAD-CO_2 emissions of the local roads reduce by nearly 2% when the eco-routing ratio increases from 10% to 50% or 100%.

An interesting observation on Figure 6.10 is that, as the budget for the local road maintenance increases, the better pavement conditions of the local roads are attracting more travellers to use them. In Figure 6.10, only PCI-based maintenance site selection is implemented. For example, at the end of Year 10 and given the IRI sensitivity factor of 3%, the AAD-VKMT on the local roads increase by 0.3%, 1.1% or 2.3% (depending on the eco-routing participation ratio) if the maintenance budget increases from 200 to 700 street blocks per year. Due to this induced demand, the higher budget for local road maintenance appears to be not so effective in reducing the CO_2 emissions for the local road network. In the most extreme case of 100% eco-routing participation rate, the increased investments in local road maintenance only lead to 0.5% reduction in CO_2 emissions in the local roads, but the savings for the whole network and the highways are 1.5% and 2.5%.



(a) Scenario 3. CO_2 emissions from the local network with different percentages of travellers choosing eco-routing.

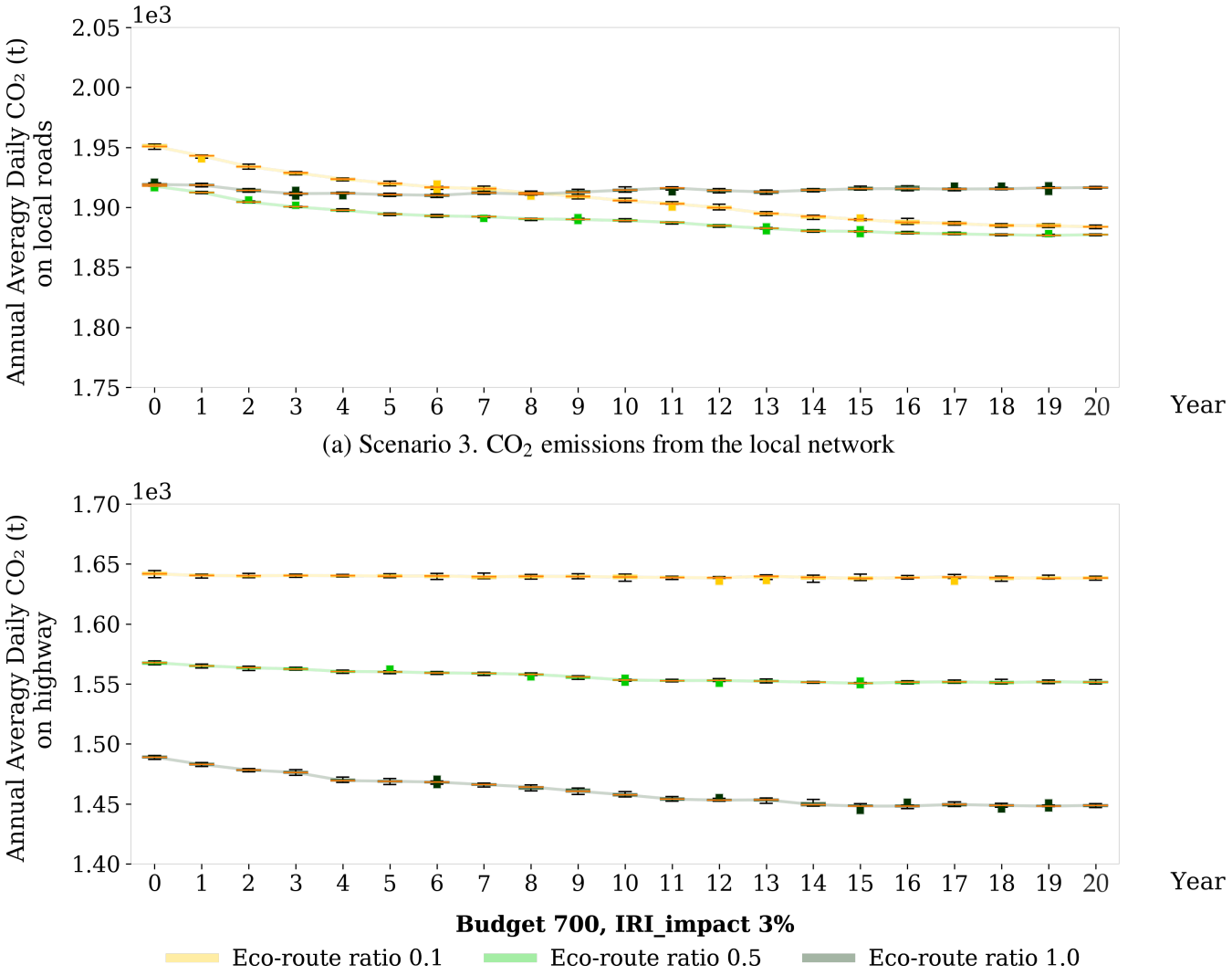


(b) Scenario 3. CO_2 emissions from the highways in SF with eco-routing strategy.

Fig. 6.10 AAD-CO_2 on local roads and highways under Scenario 3 (eco-routing).

The simulation of the eco-routing scenario involves a few random variables. First of all, travellers are randomly split into eco-routing and non eco-routing types. Secondly, as discussed in Chapter 4, the incremental trip assignment order is also random for the trips departing in the same hourly time step. To investigate whether the randomness affects the simulation results, 10 repetitions of the eco-routing scenario simulations are conducted. Only the results of subscenarios 3d, 3h and 3l are shown by the box and whisker plots in Figure 6.10 (a) and (b). The subscenario set-ups are listed in Table 6.5. For subscenarios 3d, 3h and 3l, the budget is fixed at 700 street blocks per year and the IRI impact is 3%, while the eco-routing participation ratios are 10%, 50% and 100%, respectively. The eco-routing simulations are conducted for a 20-year analysis period, as it helps to understand whether the variabilities in the outcomes change with age.

Figure 6.10 (a) displays the variability of CO_2 emissions on the local roads. The lines connect the median values of the results from 10 random repetitions. The three lines (yellow, green and grey) correspond to the results at three different eco-routing participation rates (10%, 50% and 100%). At each point (a combination of year and eco-routing ratio), a box and whisker plot with outliers are plotted. The upper and lower bounds of the box are the 25th and 75th percentiles of the 10 random repetition results ($Q1$ and $Q3$), while the whiskers extend to the most extreme value within 1.5 Inner Quartile Range ($IQR = Q3 - Q1$), namely $Q1 - 1.5\,IQR$ and $Q3 + 1.5\,IQR$. Data outside of this range are plotted as the outliers. It can be seen that the box and whiskers are very short, almost indistinguishable. The same observation is made on Figure 6.10 (b). Overall, the results of the repetition simulation with random seeds suggest that the variations in the CO_2 emission results are quite small.



(a) Scenario 3. CO_2 emissions from the local network

(b) Scenario 3. CO_2 emissions from the highways

Fig. 6.11 AAD-CO_2 on local roads and highways under Scenario 3 (eco-routing) - outcome variabilities from random experiments

Comparing eco-maintenance and eco-routing strategy

The effects of eco-maintenance and eco-routing can be compared with network-wide metrics obtained from Scenario 2 and 3. They are the two scenarios when eco-maintenance and eco-routing are implemented independently. To facilitate the comparison, Table 6.8 presents the simulation results given IRI sensitivity of 3% and maintenance budget of 700. The combined results when the two maintenance strategies are implemented together will be presented and discussed in the next section.

In terms of CO_2 emission reductions, the effect of the eco-maintenance strategy is comparable to the case when 10% trips follow the eco-routing choice. When the eco-routing participation ratio increases to 50% or 100%, it outperforms eco-maintenance in reducing CO_2 emissions.

The eco-maintenance itself does not change the traffic characteristics, but the eco-routing strategy with high traveller participation rates have negative impacts on the network efficiency. With 10% trips following the eco-routing path, the average trip duration is almost the same as the no eco-routing scenarios. When the eco-routing participation ratio increases to 50%, the average trip duration increases by 1 minute, which is still acceptable. While for 100% eco-routing, it causes an extra delay of 6 minutes. Given the journey durations are on average less than half an hour, a delay of 6 minutes is rather significant.

Besides, the eco-routing paths are on average shorter in distance than the fastest paths adopted in the eco-maintenance scenarios. Similar observations have been demonstrated elsewhere in the literature. In a study on the various routing options based on the Toyota city in Japan, the journey length for the fastest route is 9.33 km, but the eco-routing path is 3.4% shorter [252]. In another example in Boriboonsomsin et al. [29], the eco-routing distance from the Los Angeles airport to its downtown (26 km) is found to be 23% shorter than the shortest duration route (33.5 km). From the analysis in this study, it is found that on a network level, the average trip distance is around 2% or 5% less with an eco-routing participation ratio of 50% or 100%.

Finally, the eco-maintenance and eco-routing strategies are compared in terms of their influences on pavement conditions. The highways are excluded from the average PCI calculation as it is assumed that they have a relatively sufficient source of funding and are able to sustain their current conditions throughout the analysis period. When PCI-based maintenance is carried out, the average PCI of the local roads increases from 74 to 86 from Year 0 to Year 10, given the annual maintenance budget of 700 street blocks. The same improvements can be achieved when eco-routing is implemented independently. However, for the eco-maintenance strategy, as the budget and resources are diverted for the treatment of streets that are not necessarily the most deteriorated, the average pavement conditions are

lower than in the PCI-based maintenance case. The average PCI for local roads under the eco-maintenance strategy only increases slightly from 74 to 79 over the analysis period.

Table 6.8 Comparing different CO_2 mitigation measures at the network level. In the brackets, it shows the percentage or absolute change of each scenario compared to the baseline "PCI-based maintenance" (absolute changes make more sense and are used for PCI). This table is different from Table 6.9 (except for Note [2]), as this table shows the metrics at Year 0 or Year 10, while Table 6.9 shows the metrics averaged over the whole study period.

Metric	PCI-based maintenance [1]	Eco-maintenance	10% eco-routing	50% eco-routing	100% eco-routing
1. AAD-CO_2 at Year 0 (t)	3,922	3,889(-0.8%)	3,894(-0.7%)	3,795(-3.2%)	3,727(-5.0%)
2. AAD-CO_2 at Year 10 (t)	3,889	3,854(-0.9%)	3,859(-0.8%)	3,758(-3.4%)	3,688(-5.2%)
3. 10 year average daily CO_2 (t) [2]	3,905	3,869(-0.9%)	3,877(-0.7%)	3,775(-3.3%)	3,706(-5.1%)
4. Average trip duration at Year 0 (min)	24.4	24.4(-)	24.4(-)	25.6(+4.9%)	30.4(+24.6%)
5. Average trip duration at Year 10 (min)	24.4	24.4(-)	24.4(-)	25.7(+5.3%)	30.5(+25.0%)
6. Average trip distance at Year 0 (km)	9.94	9.94(-)	9.90(-0.4%)	9.74(-2.0%)	9.41(-5.3%)
7. Average trip distance at at Year 10 (km)	9.94	9.94(-)	9.90(-0.4%)	9.73(-2.1%)	9.40(-5.4%)
8. Local PCI at Year 1	77	76(+1)	77(-)	77(-)	77(-)
9. Local PCI at Year 10	86	79(-7)	86(-)	86(-)	86(-)

Note: [1] "PCI-based maintenance" is used as the baseline scenario which the outcomes from other scenarios are compared with. This row corresponds to the "network-wide (CO2) emissions" block in Table 6.9. The values are slightly different due to rounding.

6.5.3 The combined eco-maintenance and eco-routing strategies

Previous results suggest that both eco-maintenance and eco-routing can contribute to the reduction of network-wide CO_2 emissions. While each strategy also has its limitations: the effect of eco-maintenance is limited compared to eco-routing with high participation rates, while eco-routing causes traffic delays and affects the traffic operation efficiency. In this section, results will be presented and discussed for the cases when the two maintenance strategies are implemented simultaneously (Scenario 4). It aims to answer the question of whether the combined approaches can lead to more advantageous system-wide outcomes.

Figure 6.12 shows the CO_2 emissions from the local roads and highways under Scenario 4, when both eco-maintenance and eco-routing strategies are implemented. The general trends are similar to those in Scenario 3 (Figure 6.10). With an increasing eco-routing participation ratio, the CO_2 emissions from the highways clearly reduce, but not on the local roads. In fact, when the eco-routing ratio reaches 100%, due to the reduced usage of the highways, the AAD-CO_2 emissions of the local roads gradually increase towards the end of the 10 year analysis period. Nevertheless, the CO_2 emissions from the whole network are still smaller with increasing participation rates of eco-routing.

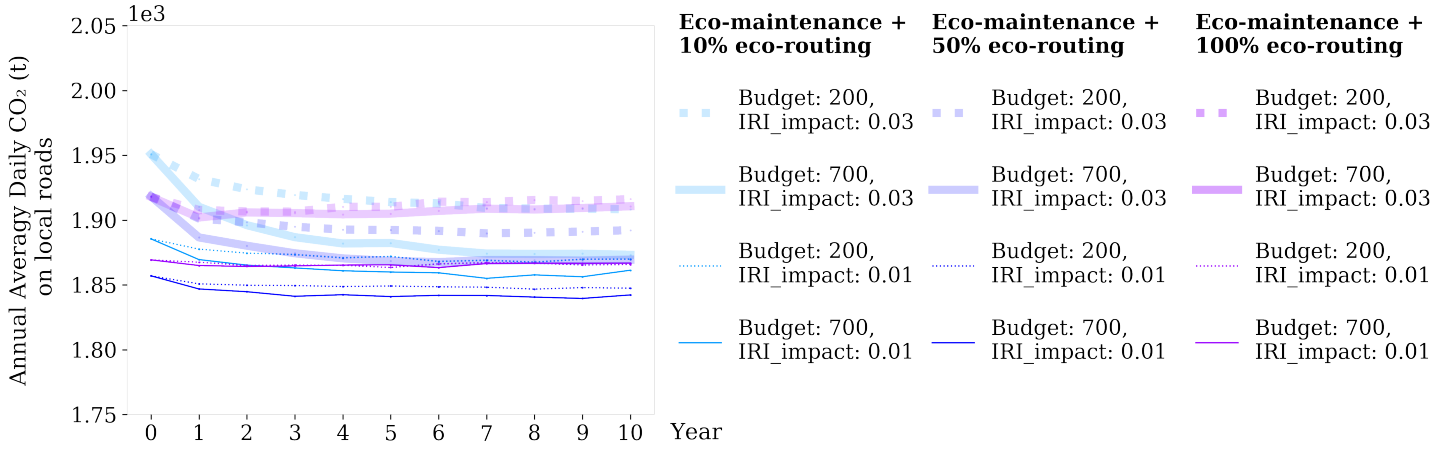
Quantitative results regarding the combined effects of eco-maintenance and eco-routing are given in Table 6.9. The results are presented for four network-level metrics. First of all, for the network-wide CO_2 emissions, the combined strategies lead to higher reductions (2% to 6%) compared to implementing eco-routing or eco-maintenance independently (1% to 5%).

Then, in terms of the traffic network efficiency, the average trip distance and duration results in Table 6.9 suggest that the combined effects of implementing the eco-routing and eco-maintenance strategies together are not different from the eco-routing only scenario. In other words, the induced traffic re-distributions and congestions due to maintenance rescheduling are not obvious at the network level.

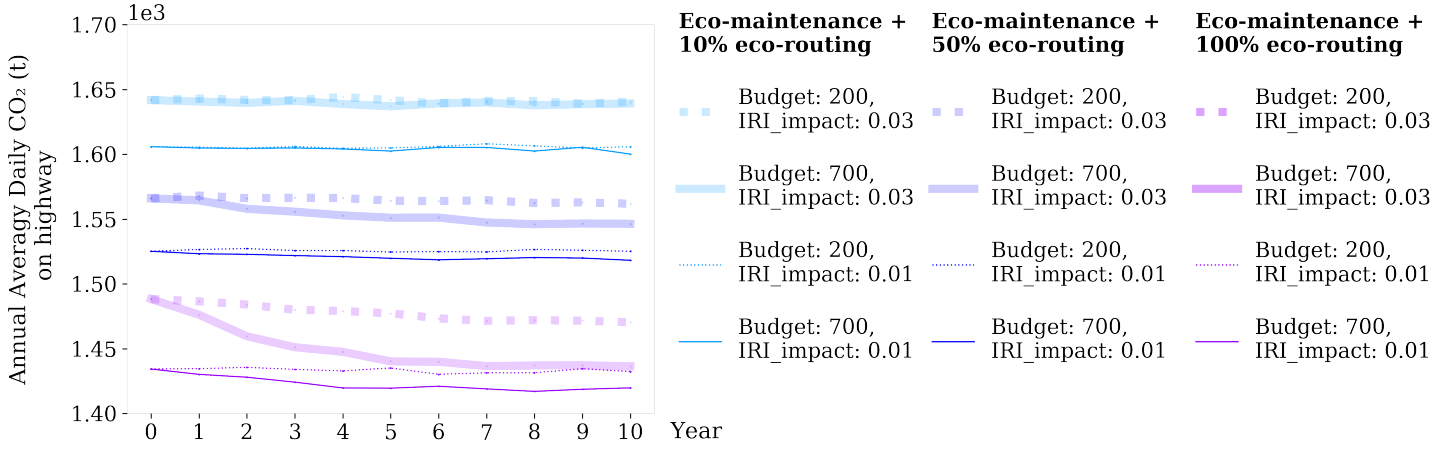
Next, as for the average pavement conditions, it is clear that eco-maintenance is not helpful if the goal is to improve the overall pavement conditions. Whenever eco-maintenance is implemented, the improvement in the average PCI is small (improves by 5 points over 10 years). When PCI-based maintenance is carried out, the 10-year improvement in the average PCI of the local roads can be as high as 12 points. This is not surprising as the conditions of pavements with low traffic volumes have just the same weight in calculating the overall PCI as the high traffic volume roads. The average PCI may be an important indicator in terms of social equality, as people living in less populated neighbourhoods deserve to receive a fair amounts of infrastructure investments as the residents in the downtown area. Regrettably,

this equality concern has not be incorporated in the design of the eco-maintenance strategy in this study.

More detailed results regarding the network-wide metrics under different sensitivity scenarios are given in Appendix 2.



(a) Scenario 4. CO_2 emissions from the local network under both eco-maintenance and eco-routing strategies



(b) Scenario 4. CO_2 emissions from the highways in SF under both eco-maintenance and eco-routing strategies

Fig. 6.12 AAD-CO_2 on local roads and highways under Scenario 3 (eco-routing) and 4 (eco-maintenance and eco-routing).

6.6 Additional scenarios considering traffic growth

The analysis in the previous sections in this chapter does not consider factors such as traffic growth, fuel mix change or climate change. The reasons for excluding these factors are two fold. First of all, their future trends are hard to predict and their impacts would ideally be handled with extensive sensitivity analysis. Secondly, it is to separate the impacts of pavement management and traffic operation on CO_2 emissions, traffic efficiency and infrastructure conditions from those caused by the increasing levels of traffic and other factors. However, in order to have a basic understanding of the implications of the assumptions in this chapter, an additional set of scenarios is presented in this section, incorporating the forecast traffic growth into the simulations.

6.6.1 Source of traffic growth information

The source of the traffic growth forecast comes from the ConnectSF project, a multi-agency collaborative transportation plan for the city of SF [54]. Key government agencies in transportation planning and management are involved, including the SFCTA, SFMTA, San Francisco Planning Department and others. As of 2019, ConnectSF has just embarked on its second phase, which serves to (1) identify the existing and future travel needs and options in SF, and to (2) develop major projects to meet these needs. To facilitate reaching these goals, ConnectSF conducted a series of studies, releasing the outcomes of them such as the zone-level population, employment, trip pattern and travel time changes from 2015 to 2050. Specifically, the trip pattern data from ConnectSF are used to infer the vehicular traffic growth rate for the analysis in this study [55]. In the dataset, the regions in and around the city of SF are divided into 15 districts (12 districts within SF and 3 districts representing the areas to the south, east and north of SF), as shown in Figure 6.13. For each district, numbers are provided regarding the total arrivals and departures on a typical day by various transport modes (automobile, transit, walking, etc.). Only trips using driving-related modes are retained for the study in this section. After determining the total number of driving trips originating/ending in each district, Equations 6.5a and 6.5b are used to calculate the annualised traffic growth rates for each zone. As no data in the intermediate years between 2015 and 2050 is provided, the annualised traffic growth rates are calculated based on the trip counts of 2015 and 2050. The growth rates for incoming and outgoing trips are calculated separately according to the numbers of trips arriving and departing from each

Table 6.9 The combined effects of eco-maintenance and eco-routing, assuming IRI sensitivity of 0.03 and a budget to repair 700 street blocks per year

1. Network-wide CO_2 emissions		
Baseline average daily CO_2 emissions (t) over the analysis period in Scenario 1 (PCI-based maintenance and shortest-time routing)		
Baseline value	3,905	
Eco-maintenance	3,869	-1%
10% Eco-routing	3,877	-1%
50% Eco-routing	3,775	-3%
100% Eco-routing	3,706	-5%
Combined 10% eco-routing	3,841	-2%
Combined 50% eco-routing	3,744	-4%
Combined 100% eco-routing	3,677	-6%
2. System efficiency in terms of VHT		
Baseline average trip duration (min) over the analysis period in Scenario 1		
Baseline value	24.4	
Eco-maintenance	24.4	-
10% Eco-routing	24.4	-
50% Eco-routing	25.6	5%
100% Eco-routing	30.6	25%
Combined 10% eco-routing	24.4	-
Combined 50% eco-routing	25.6	5%
Combined 100% eco-routing	30.9	27%
3. System efficiency in terms of VKMT		
Baseline average trip distance (km) over the analysis period in Scenario 1		
Baseline value	9.94	
Eco-maintenance	9.94	-
10% Eco-routing	9.90	-
50% Eco-routing	9.73	-2%
100% Eco-routing	9.41	-5%
Combined 10% eco-routing	9.90	-
Combined 50% eco-routing	9.73	-2%
Combined 100% eco-routing	9.40	-5%
4. Average PCI of local roads, Year 0 to Year 10		
Baseline PCI improvements over the analysis period in Scenario 1		
Baseline value	74 to 86	+12
Eco-maintenance	74 to 79	+5
10% Eco-routing	74 to 86	+12
50% Eco-routing	74 to 86	+12
100% Eco-routing	74 to 86	+12
Combined 10% eco-routing	74 to 79	+5
Combined 50% eco-routing	74 to 79	+5
Combined 100% eco-routing	74 to 77	+3

district. However, the two growth rates are very similar for all districts, thus presented as a single number in Table 6.11.

$$O_{i,2015} \times (1 + g_{i,O})^{2050-2015} = O_{i,2050} \tag{6.5a}$$

$$D_{i,2015} \times (1 + g_{i,D})^{2050-2015} = D_{i,2050} \tag{6.5b}$$

where $O_{i,2015}$ and $O_{i,2050}$ are the numbers of trips originating from District i in 2015 and 2050; $D_{i,2015}$ and $D_{i,2050}$ are the numbers of trips ending in District i in 2015 and 2050; $g_{i,O}$ and $g_{i,D}$ are the annual traffic growth rates for trips originating from or ending in District i.

As seen in Table 6.11, the growth of car traffic in SF is relatively slow, with the growth rates in most districts being less than 1% per year. This could be because of the already saturated traffic conditions in the city and not so much room has been left for further growth. After obtaining the district-level trip growth rates, Traffic Analysis Zone (TAZ) level OD pairs are generated for each year in the decade long study period. These new OD pairs reflecting traffic growth are then used to calculate the network-wide traffic efficiency, infrastructure condition and emission metrics. Numbers of trips on Wednesday, the representative day of week used for the analysis, are shown in Figure 6.14. The total daily traffic at the end of the analysis period is 9% higher than in the initial year. Over the whole analysis period, the average daily traffic is 4.6% higher than in the initial year, or equivalently 4.6% higher than not considering any traffic growth.

6.6.2 Results considering traffic growth

The same 32 scenarios shown in Table 6.5 are simulated again, but this time considering traffic growth. For the purpose of clarity, only the results from 16 scenarios, namely those with an IRI impact factor of 3%, are presented below.

As there are more cars on the road, a direct consequence of considering traffic growth is the increase of AAD-VHT and AAD-VKMT, shown in Figures 6.15 and 6.16. The results presented in these two figures are divided into eight groups, each referring to a specific maintenance and routing scenario (shown on the vertical axes of the plots). Each group contains four horizontal bars, each representing two budget levels (the unhatched bars: 700 blocks of roads maintained each year; the hatched bars: 200 blocks of roads maintained each year) and two traffic growth assumptions (e.g., grey: no traffic growth; red: with traffic growth for the AAD-VHT plot). Each horizontal bar consists of two parts: the lighter shade rectangle on the left represents the AAD-VHT (or AAD-VKMT) of the local roads, while the darker shade rectangle on the right shows that of the highways. A small number of privately

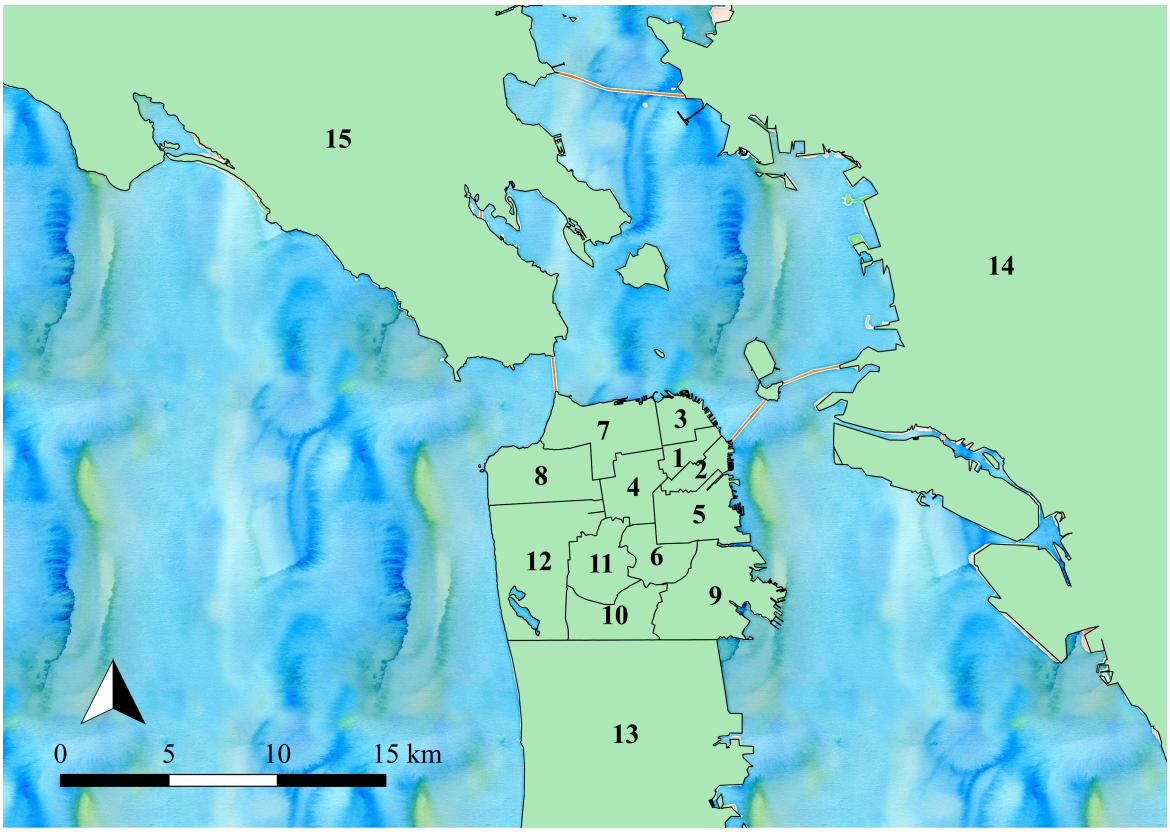


Fig. 6.13 Districts in ConnectSF trip pattern data.

Table 6.11 Names and traffic growth rates for ConnectSF districts

No.	Name	Traffic growth rate	No.	Name	Traffic growth rate
1	Downtown	0.009	**9**	Bayshore	0.016
2	South of Market	0.023	**10**	Outer Mission	0.009
3	North Beach/ Chinatown	0.006	**11**	Hill Districts	0.007
4	Western Market	0.006	**12**	Sunset	0.006
5	Mission/Potrero	0.014	**13**	South Bay	0.006
6	Noe Valley/ Glen Park/ Bernal Heights	0.004	**14**	East Bay	0.006
7	Marina/ North Heights	0.004	**15**	North Bay	0.005
8	Richmond	0.005			

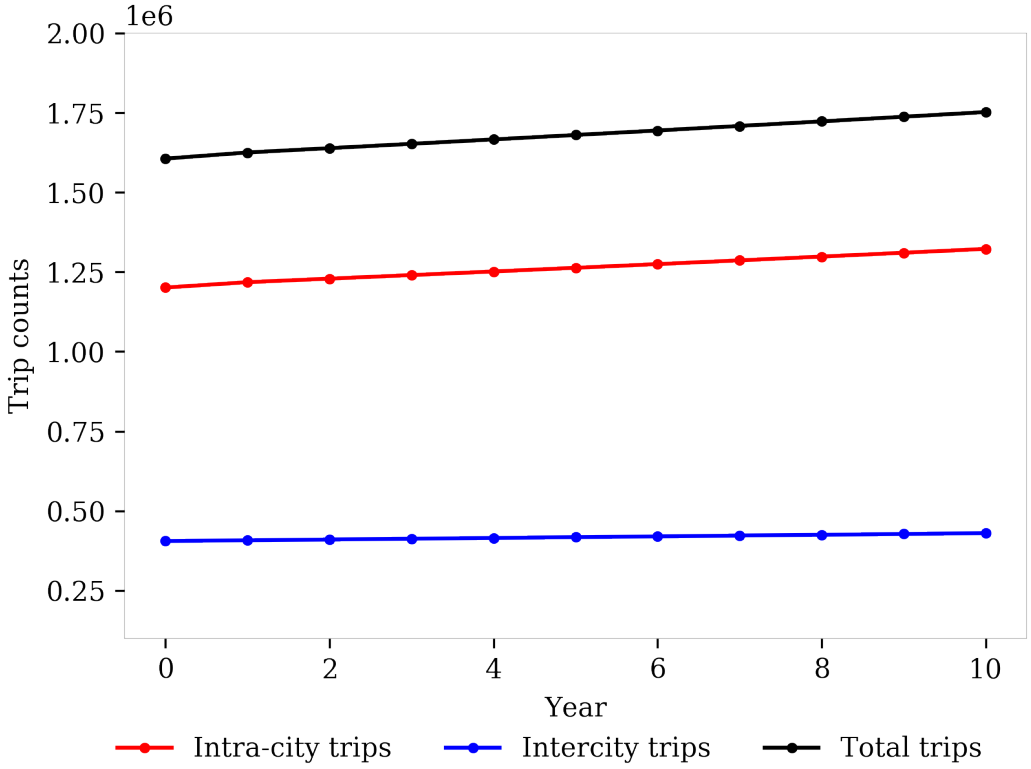


Fig. 6.14 Trip counts on a representative day of week throughout the analysis period.

managed roads are not shown in the figures. According to the results presented in Figure 6.16, due to traffic growth, the AAD-VKMT increases by about 4.8% to 5.5% on the local roads, about 3% to 3.7% on the highways, or around 4.3% if looking at the whole road network. These figures are in agreement with the 4.6% increase in the 10-year average trip counts when considering traffic growth. As for the AAD-VHT, it is expected that the travel time grows superlinearly with the traffic volume, as the volume-delay relationship (Equation 3.3 in Section 3.5.2) indicates a time delay factor to the power of four with any increase in the traffic volume. Indeed as shown in Figure 6.15, the vehicle travel time increases by 6.7% to 8.5% on the local roads, 5.8% to 7.2% on the highways and 6.7% to 9.5% for the whole network when considering traffic increase. The increase in AAD-VHT is larger than that of the AAD-VKMT, indicating a longer travel time for unit distance (or a lower speed of the vehicles) if traffic increases as predicted.

Figure 6.17 compares the 10-year average AAD-CO_2 under different budget, maintenance and routing scenarios with and without traffic growth. As expected, the AAD-CO_2 on the local roads and highways are consistently higher when considering the traffic growth,

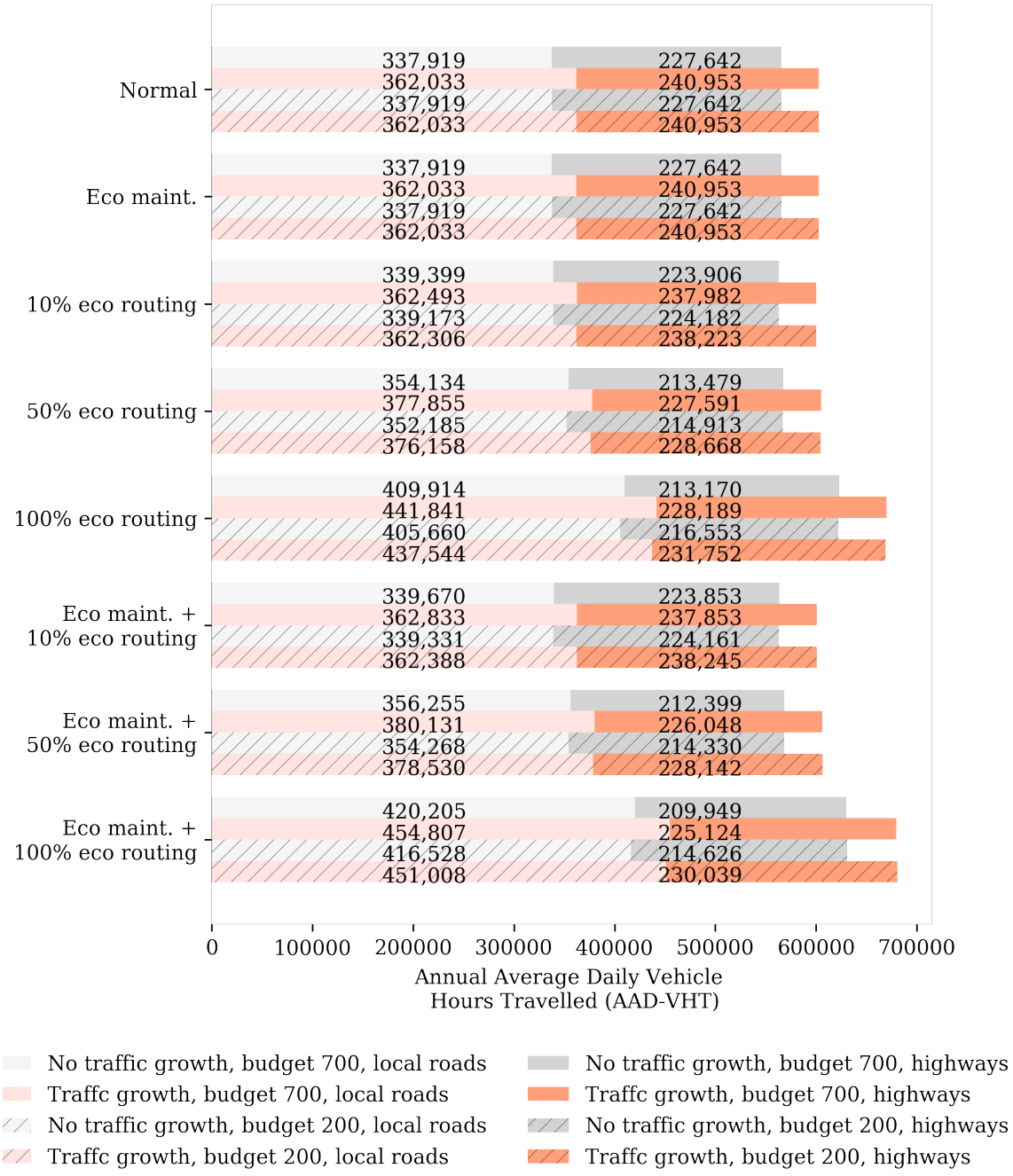


Fig. 6.15 AAD-VHT on local and highways road network with or without traffic growth.

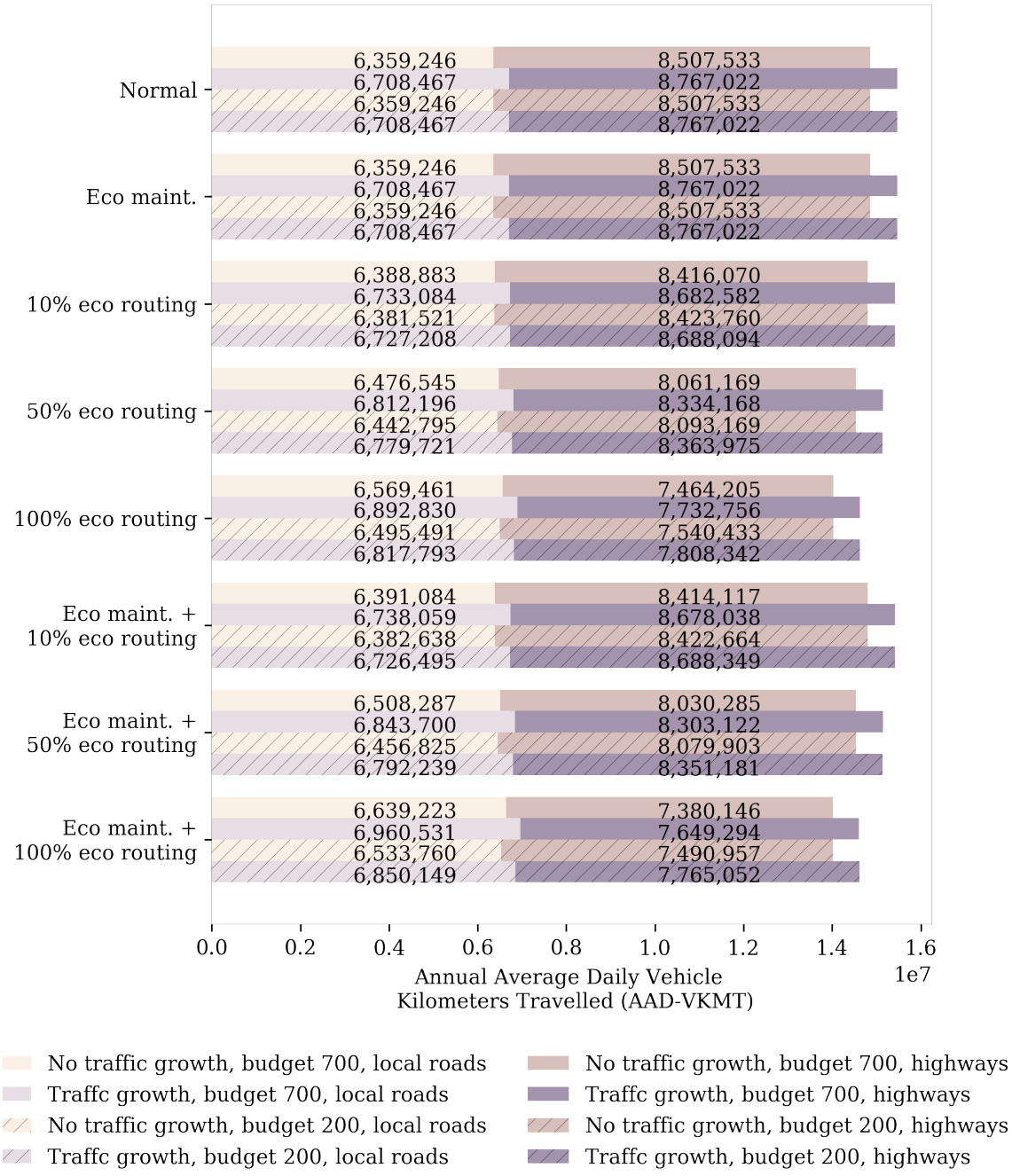


Fig. 6.16 AAD-VKMT on local and highways road network with or without traffic growth.

specifically around 6% higher for the local road network and around 4% higher for the highways. For the whole road network in SF (consisting of the local roads, highways and privately managed roads), the increase in AAD-CO_2 after considering traffic growth is around 5.4%. It should be noted that the magnitude of increase in AAD-CO_2 due to traffic growth alone (not considering cleaner fuel, etc.) is so large that only 100% eco-routing can possibly offset this increase. However, according to the vision of ConnectSF, despite the increase in traffic and reduction in vehicle running speed (to 15 mph in 2050), the GHG emissions are still predicted to be 26% lower by 2050 compared with the 2015 case, thanks to fuel efficient technologies and the use of low carbon traffic modes such as cycling and public transit [224]. While not incorporated in this thesis, these factors can be incorporated into future studies to investigate the elasticity of the outcomes given diverse future scenarios.

Lastly, the average pavement conditions of the local roads at the end of Year 10 under various budget, maintenance, routing and traffic growth scenarios are presented in Figure 6.18. From this figure, it can be seen that the key factors affecting the pavement conditions are the maintenance strategy and maintenance budget. The routing strategy and traffic growth have almost no impact on the average pavement conditions. It is a likely consequence as the traffic load is not found to be a significant factor related to pavement degradation. Further analysis into the streets picked for maintenance indicates that, in the initial year, the roads selected for maintenance are 100% the same in both of the traffic increase and no traffic growth case. This is certainly due to that the travel demand and traffic distributions are the same for the initial year in both cases. While such similarity gradually decreases and at the end of year 10, only 64% overlap exists between the roads selected for maintenance with or without considering traffic growth.

6.7 Discussions

The magnitude of CO_2 emissions and reductions

According to a report by the SF Department of Environment, the total GHG emissions from passenger vehicles in SF in 2017 was 1.67 million metric tons of CO_2 equivalent (consisting of CO_2, methane and nitrous oxide, etc.) [211], with which the estimation in this study of 3,900 tonnes of CO_2 per day is in general in agreement. However, the magnitudes of CO_2 emission reduction after implementing eco-maintenance and eco-routing appear to be small in this study. For eco-maintenance, the cumulative savings in CO_2 emissions are only 0.6% to 1.9%, depending on the budget level and IRI sensitivity factor. While for a realistic eco-routing participation rate of 10%, the 10 year cumulative savings in CO_2 emissions

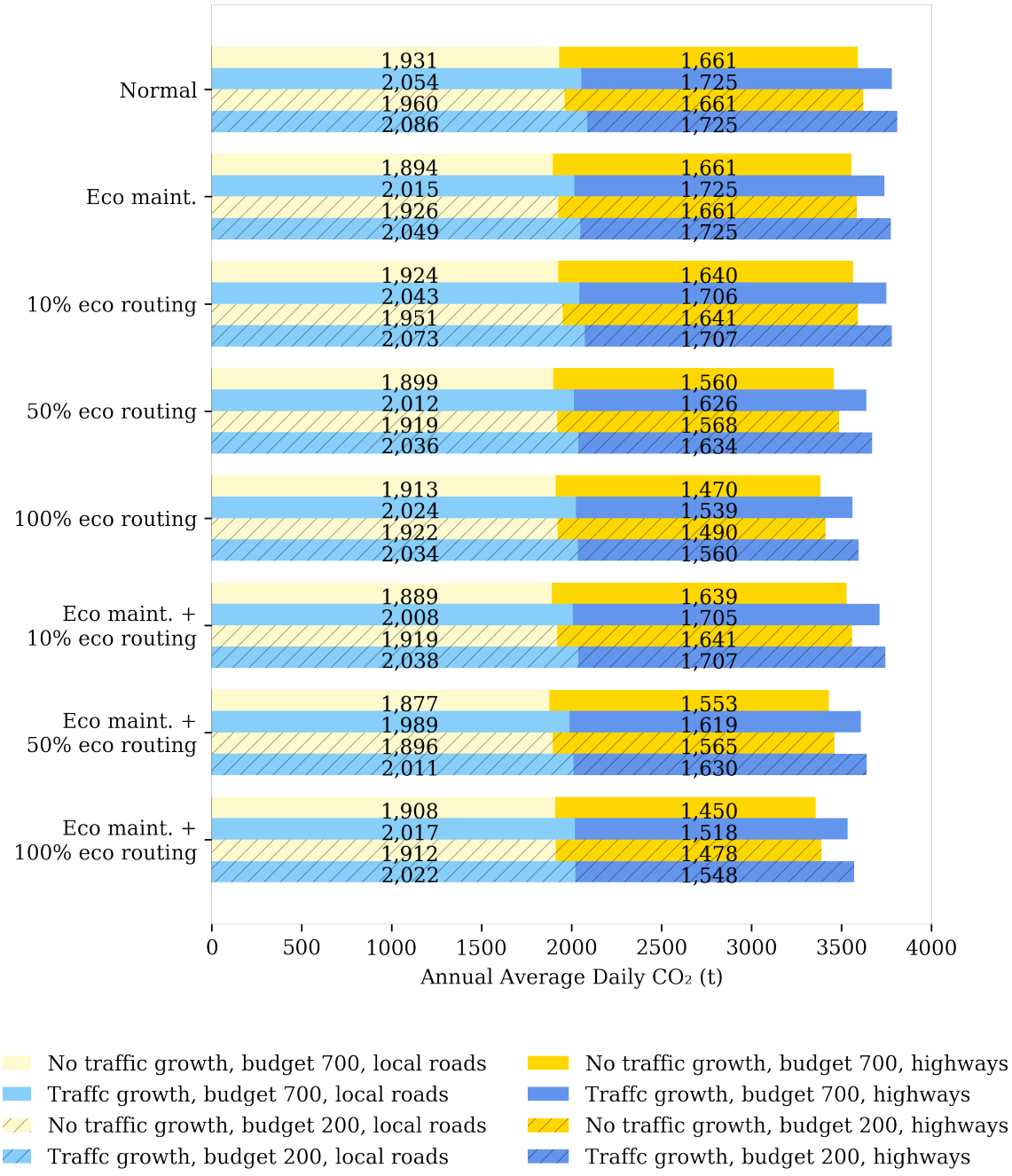


Fig. 6.17 AAD-CO_2 on local and highways road network with or without traffic growth.

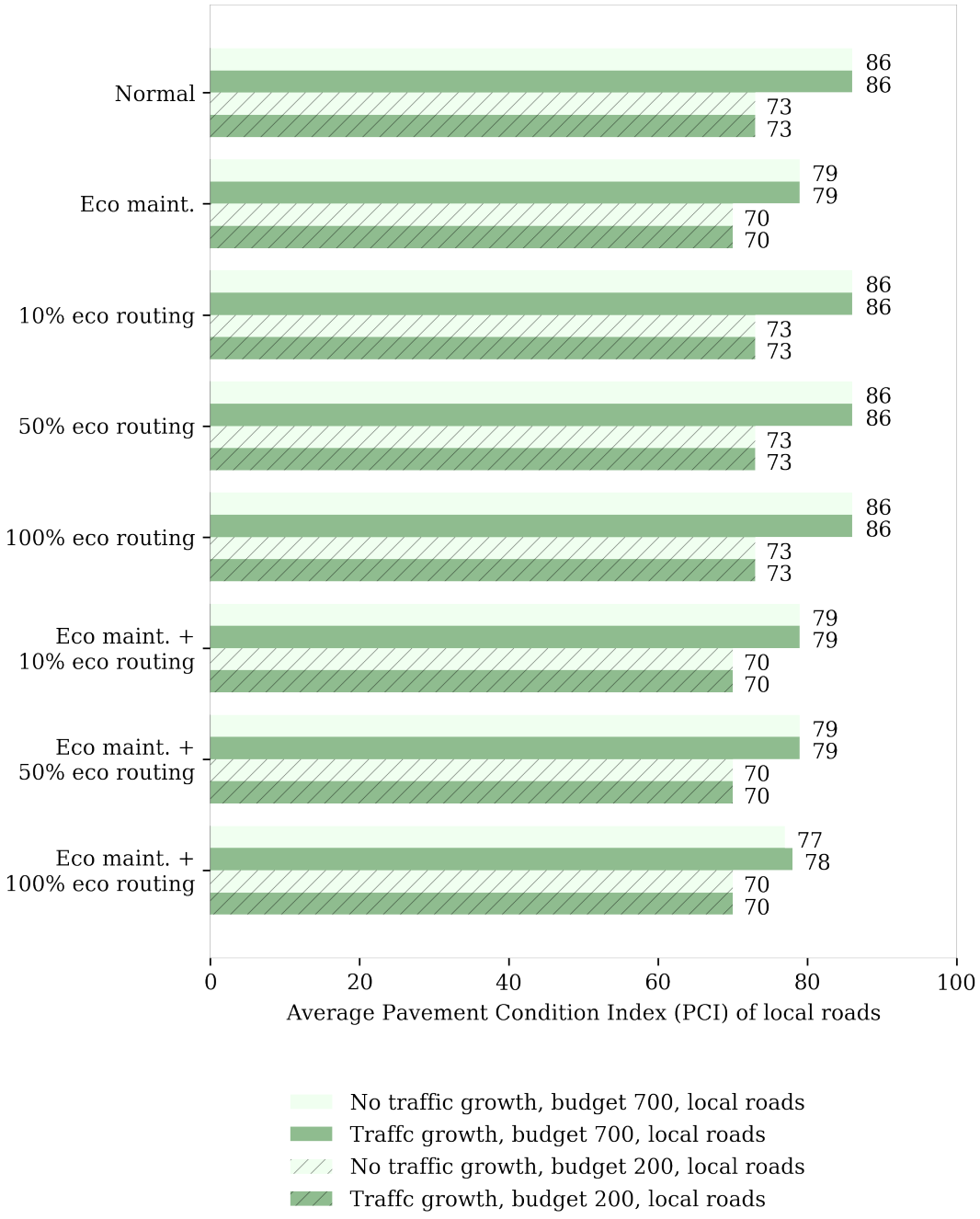


Fig. 6.18 Average PCI of local roads with or without traffic growth.

are also only 0.7% (it reaches 5% with 100% eco-routing participation rate). To check the sensibility of the conclusions, these results are compared with existing literatures.

An eco-friendly path searching system is proposed in Zeng et al. [252]. By studying nearly 8,000 trips (origin-destination pairs), the authors are able to construct an empirical relationship between CO_2 emission savings of the eco-routing with respect to journey distance and time budget. It is found that for journeys with distance around 10 km (as in the SF case study in this thesis), 0-6% of savings in CO_2 emissions are expected when switching from the fastest to the least emission routes. Another study reports that for trips with distance less than 15 miles (24 km), the fuel savings with eco-routing are around 12% on average compared with the fastest routes [29]. In Ahn and Rakha [1], 3 to 5% reductions in CO_2 emissions are reported for the eco-routing system at the network level based on dynamic User Equilibrium (UE) traffic simulations on two real networks. According to these results, the CO_2 reduction of a maximum of 5% obtained in this chapter appears to be reasonable.

The effect of eco-maintenance is considerably weaker compared to eco-routing in this thesis. This is expected because only the maintenance scheduling on the local road network is studied in the analysis. In Wang et al. [240], life cycle analysis is carried out on four highway segments. The reductions in the use phase emissions of the road with the lowest traffic level (AADT of 3,200) are only 1% of that of the busiest road (AADT of 86,000). Similar ideas are also suggested by Santero et al. [203], where an extra rehabilitation is said to be a potentially cost-effective method for reducing emissions on roads with high traffic volumes.

The effect of combined eco-maintenance and eco-routing strategies

Results in Table 6.9 show that the combined effects of eco-maintenance and eco-routing in terms of several network-wide metrics are not different from the simple sum of those when the two emission mitigation strategies are implemented independently. On the one hand, this suggests the lack of strong interactions between the traffic simulation and the pavement degradation modules. Indeed, many interesting interactions are not considered due to the computational difficulty or lack of evidence. For example, traffic delays due to roadworks are not included due to the speed of the current traffic simulation module. While the traffic load is found to be not a good predictor of pavement degradation given the specific pavement degradation dataset.

On the other hand, the weakly coupled traffic simulation and pavement degradation models are still able to capture certain interactions. For example, the travellers being "attracted" to use the local roads when local road conditions are improved after implementing eco-maintenance or with a higher budget, as well as the change in the average pavement condi-

tions in the long term after introducing eco-routing. However, these localised "compounding" effects between eco-maintenance and eco-routing diminish when using the network-level summary metrics.

Chapter 7

Conclusions and recommendations for further work

This thesis presents a study on the effects of various carbon mitigation strategies evaluated based on the CO_2 emission, traffic efficiency and infrastructure condition metrics for a city-scale transportation network. The innovation as well as the focus here is to adopt an interdisciplinary perspective that includes not only the traffic operational aspect, but also the transport infrastructure management knowledge. The city of San Francisco (SF) is chosen as the background for this study due to the availability of a variety of open data in this area. As a conclusion of the thesis, this chapter first revisits the research objectives set in Chapter 1 and discusses about how they have been met. A detailed summary of the findings from each chapter is presented in Section 7.2. This is followed by a discussion on the policy implications in Section 7.3 and recommendations for future studies in Section 7.4.

7.1 Revisiting the research objectives

Two research objectives have been set for this research in Section 1.2, including the development of the city-scale traffic and pavement degradation modelling tools, as well as the deployment of these two modules jointly for carbon mitigation scenario analysis. In terms of tool development, a mesoscopic traffic simulation model is presented in Chapter 3. It is shown to be able to run efficiently with a city-scale network and produce hourly varying traffic patterns. Preliminary validations against an existing model show good agreements between them. The validity of the traffic simulation is further investigated in Chapter 4. It is found that the assumption of "perfect road condition information" being available to every road user does not lead to noticeable different results, compared with those obtained

under more realistic assumptions of imperfect, bounded knowledge of the system. Chapter 5 presents the development of the pavement degradation model, which adopts an innovative spatial structure to overcome the drawbacks in the available data, such as missing or erroneous records. A sub-objective of the tool development is to identify the connections between the traffic simulation and the pavement degradation modules. For example, pavement maintenance will lead to traffic disruptions, while changes in traffic patterns will also alter the use phase CO_2 emissions associated with each pavement segment. However, in Chapter 5, traffic load is found to perform worse in predicting pavement degradation compared with the age-based model, given this specific dataset used in the study. As a result, the impacts of traffic load on pavement degradation are not considered.

Based on the city-scale modelling tools developed, the second objective involves the evaluation and comparison of the impacts on the transportation system-wide CO_2 emissions, traffic efficiency and infrastructure condition metrics brought by different carbon mitigation scenarios. This procedure is established in Chapter 6, where the quantification of the network-wide metrics (especially the CO_2 emissions) is achieved by considering the influences of individual road gradient, speed, pavement roughness in a step-wise manner before taking a summation over the whole network. Different carbon mitigation scenarios are then represented by changing the corresponding set-ups in the joint simulation, such as the driver's route choice behaviour (from time-based routing to eco-routing) or maintenance site selection priorities (from condition-based to CO_2 reduction oriented eco-maintenance). The evaluation framework is shown to perform well and the results exhibit reasonable sensitivities to traffic growth, changes in traffic operations and pavement maintenance practice.

7.2 Summary of the main findings

An introduction of the research topic and the background of transportation system carbon mitigations are given in Chapter 1. Data suggest that the CO_2 emissions produced by the road traffic are the most significant component in the whole transport sector, while the total CO_2 emissions from the road transport sector change over time as the traffic demand increases and the vehicle fuel economy improves. It is identified that the Intelligent Transportation Systems (ITS) offers promising new opportunities in managing the vehicle traffic CO_2 emissions in the digital age. Also, it is receiving increasing awareness among the engineers that the pavement conditions may have important impacts on the fuel consumption and CO_2 emissions of the vehicles. In fact, the use phase CO_2 emissions are the dominant component in the Life Cycle Assessment (LCA) of the pavement infrastructure, especially for high traffic volume roads. Despite the great opportunities in CO_2 mitigation from both the traffic operation and

pavement maintenance perspectives, there are hardly any existing research that combines and integrates these two aspects to assess the carbon emissions of a transport system.

To facilitate the understanding and integration of the traffic operation and pavement maintenance scenarios, existing theories and tools related to the traffic and pavement condition performance modelling are reviewed in Chapter 2. It begins with a review of the sustainability and sustainable transport concept. Although this study is not designed to be a comprehensive evaluation of the transport sustainability, it still benefits from the many examples and indicators developed in the sustainability literature. The second part of Chapter 2 focuses on the traffic modelling literature, which is usually regarded as the feedback interactions between the road network supply and the trip-level or individual/household-level travel demand. Trip-based travel demand models are less flexible in reflecting individual person-level activity schedules and re-scheduling. However, they are less complex (easier to develop) and can take advantage of the real world data from many existing travel demand surveys. In terms of the network supply simulations, vehicle flows can be simulated by microscopic, mesoscopic and macroscopic flow propagation models. The speed of the macroscopic simulations do not depend on the numbers of vehicles, thus making them scale better to large networks. Microscopic models offer highly detailed representations of individual vehicles' movements, but are more time consuming to run, especially for large scale simulations. The mesoscopic models are in between of the macroscopic and microscopic approaches in terms of complexity and fidelity. Through the selection of the desired model features (e.g., the efficient macroscopic interaction rules and the disaggregated microscopic behaviours), mesoscopic models are deemed suitable for large-scale simulations with some person-level behaviour considerations.

In the last section of Chapter 2, existing pavement degradation prediction models are reviewed, with particular focus on their adopted methodologies, inputs data, output parameters (different quantifications of the pavement degradation) and applicability. Like the traffic models, the high fidelity pavement degradation models based on mechanistic principles are data and computationally intensive, while empirical or probabilistic models based on statistical analysis are more preferable for city-scale analysis or for planning purposes. Pavement age and the cumulative traffic load are the two most widely used parameters for degradation prediction. While degradations are quantified by individual types of defects (rutting, cracking, potholing, ravelling, edge breaking, etc.) or a compound index in other cases. Existing pavement degradation models tend to produce a single or a family of pavement degradation prediction functions for a representative analysis unit, and the proposed degradation function may not generalise well to applications outside of the study area.

Mesoscopic traffic modelling

Based on the literature review, a trip-based mesoscopic traffic simulation model is developed. It combines the macroscopic Bureau of Public Roads (BPR) volume-delay relationship [34] with disaggregated travel demand. The travel demand inputs are informed by the origin, destination and journey time information of the Traffic Network Companies (TNCs) obtained by a previous study [213]. The traffic distribution patterns are simulated by assigning trips to their fastest routes based on an incremental Static Traffic Assignment (STA) procedure. Within each hourly time step, the trip assignments occur in 20 sub-steps. While within each sub-step, the traffic conditions are assumed to be constant. This updating interval (each sub-step) is the fundamental justification that permits the parallelisation of the time-consuming trip assignment computation. As the main purpose of this chapter is to document the development of the traffic simulation module, the findings from this chapter are mostly related to the performance of the proposed tool: The proposed model can capture the time-varying traffic distributions across a city-scale network. Specifically, the traffic simulation model can produce traffic volume and speed for every drivable road links in SF (about 27,000 road links) at a temporal resolution of one hour. With parallelisation and adopting the efficient priority-queue based Dijkstra's shortest path algorithm, it takes around 40 minutes to simulate the route choices of 9 million travellers for a typical week.

The mesoscopic traffic simulation presented in Chapter 3 assumes that all travellers always have access to the real-time traffic condition information at the time of departure. The plausibility of this assumption is investigated in Chapter 4 by revealing the traffic conditions of parts of the road network. In such a scenario, only certain "probe" vehicles can collect and share their experienced travel time in each simulation sub-step. Thus, for the road links not used by the probe vehicles in the current sub-step, no data or only outdated information exists regarding their current congestion states. This probe vehicle scenario is further extended by varying the probe penetration ratio (percentages of the probes among all vehicles) and the probe information variability COV (coefficient of variation of time delay per unit distance). A COV of 1 indicates that the mean and the standard deviation of the unit time delay are equal, and a COV of 1 or less has been observed in the real life. It is found in this study that given a probe information variability COV of 1 and a probe ratio of 1%, the average trip travel time from the traffic simulation is 2 minutes longer than the perfect information case (COV = 0 and probe ratio = 100%). In other words, the traffic efficiency (measured by the average trip travel time) is not compromised when the probe ratio is $\geq 1\%$ and COV is ≤ 1. This condition should be realistic for the case study area of SF, where the probe penetration rate (e.g., smart phone navigation application users) is quite high. Two other findings from this chapter also relate to the general assumptions of the traffic simulation model. First of all,

the variabilities of the traffic simulation outcomes (average trip travel time by hour) are not big due to the random splitting of trips into sub-steps, according to the variability analysis of 10 repeated experiments. The randomness in the trip assignment order does not usually make a difference in the average trip travel time of more than 1 minute. Moreover, the average trip travel time is less than 40 minutes even during the most congested evening peak (6 PM on Friday), which justifies the assumption of the static traffic assignment.

The mesoscopic traffic simulation module produces spatio-temporally varying traffic patterns, which feed back to the route calculation as the journey time costs to the next batch of trips. Various measures of the traffic operation efficiency can be obtained based on the results of the traffic simulations, such as the Vehicle Hours Travelled (VHT), Vehicle Kilometres Travelled (VKMT), link-level delays, volume-to-capacity ratios, etc.

City-scale pavement degradation analysis

In Chapter 5, a city-scale spatial pavement degradation model is built upon 20 years of pavement condition observation surveys for the case study city. The survey reports the degradations of individual street segments in terms of a compound index, the Pavement Condition Index (PCI). Other factors that are available from the survey data include the maintenance dates, a brief description of the material types and functional classes. The dataset is augmented in this study by linking with the traffic simulation results. The passenger vehicle volumes from the mesoscopic traffic simulations and the bus volumes based on the General Transit Feed Specification (GTFS) are converted to the cumulative Equivalent Single Axial Load (ESAL), a measure of the pavement damage potential of the traffic loading. Pavement age since last maintenance and the cumulative traffic loading ESAL are the two choices of the degradation predictors. In addition, three pavement degradation model forms are also compared. With the training and testing errors both being the largest (Root Mean Square Error, RMSE, larger than 11), the non-spatial models that divide the pavements into 14 categories (by material types, road classes and bus routes) are found to be insufficient in capturing the degradation trend. The model performance can be improved by treating each street itself as a category, i.e., proposing a degradation relationship for every street segment in the dataset. However, such a model suffers from the data availability and quality issues, as there are not enough observations per street in the input dataset to justify an individual street-based model. The last degradation model form tested is a spatial-explicit structure, where the degradation rates of neighbouring pavement segments are assumed to correlate. The spatial model leads to a smaller RMSE (7 to 10) than the non-spatial categorical model. Compared with the non-spatial individual street-based model, the spatial model has similar

performance in the data fitting, but nevertheless it is able to compensate for the missing data and lead to an estimation of the degradation rates for all the street segments.

Chapter 5 is important in the overall methodological framework of system-wide carbon emission analysis, not only for the development of the pavement infrastructure performance model, but also for determining the interactions between the traffic operation and pavement degradation. Unfortunately, given various limitations of the data, the effect of the cumulative traffic load is found to be not as strong as the age effect in modelling pavement degradations. In fact, when using the log cumulative ESAL as the degradation predictor, the RMSE of the model is always 1 point (1 PCI) higher than the age-based model. One reason behind this lack of interaction between traffic and pavement degradation is due to the unavailability of the actual traffic loading history for the case study area. The vehicle volume data based on the mesoscopic traffic simulation model may not accurately reflect the real vehicle load experienced by the pavements. However, another issue that is realised through the modelling process is that the visual survey based pavement condition observations contain big variabilities and many inconsistencies. As a result, even with the accurate traffic information, it is unclear if the model fitting will improve significantly.

The pavement degradation model outputs individualised degradation forecasts for all the streets in the city. Pavement conditions are set to a higher value whenever the maintenance activities are carried out to reflect the improvements in pavement conditions due to the maintenance work.

Transport system-wide carbon emissions, traffic efficiency and infrastructure conditions

The traffic simulation model and the pavement degradation model are combined together into a system-wide analysis tool in Chapter 6 for the simulation of various carbon mitigation scenarios. Vehicle CO_2 emissions can be calculate as the multiplication of the travel distance and the unit distance CO_2 emissions. The unit distance CO_2 emissions again depend on the road gradient, traffic condition (average speed) and the pavement condition (surface roughness). The system-wide CO_2 emission simulation is related to the traffic simulation and the pavement degradation modules through two loops. In the first loop, the traffic state outputs (average link-level speeds) from the traffic simulation module are used directly in the calculation of the street-level CO_2 emissions. While the street-level emissions feed back to the traffic simulation as some travellers choose the routes that can minimise the CO_2 costs of their trips (eco-routing). In the second loop, the pavement degradation and maintenance effects are taken into consideration. Pavement roughness increases annually due to the natural degradation process. The change in pavement roughness directly affects

the road-level CO_2 emissions. Under the eco-maintenance scenario, where the maintenance site selections prioritise roads with greater emission reduction potentials, the resulting traffic distributions based on the traffic simulation module affect the decisions of where to carry out roadworks.

Two transportation system-wide CO_2 emission mitigation strategies are tested in Chapter 6. They are the eco-maintenance strategy from the pavement asset management perspective, as well as the eco-routing strategy from the traffic operation side. Quantitative results and qualitative observations regarding the independent and combined effects of these two strategies are listed below. More detailed results are provided in Appendix B. The majority of the results are conducted without considering traffic growth (to separate the influence of traffic growth from those of traffic operations and pavement maintenance). When traffic growth is considered, the changes in the key results are given at the end of this section.

(a) The eco-maintenance strategy:

Maintenance site selections are only carried out for the local road networks due to the availability of data as well as to fill in the research gap. Highways are assumed to be in constant good conditions. Streets selected for maintenance under the eco-maintenance strategy are different from those selected under the PCI-based maintenance strategy. In the eco-maintenance case, some high traffic volume streets are repaired more often, as any degradation on these streets will cause bigger increase in system-wide CO_2 emissions.

The CO_2 emission savings from the local road network under the eco-maintenance scenario depend on the budget level and the IRI sensitivity. The cumulative CO_2 emission savings over a 10-year analysis period are 1.7% to 1.9% compared to the PCI-based maintenance case given an IRI sensitivity factor of 3% (1.7% for low budget level and 1.9% for high budget level). While the savings are only 0.6% if the IRI sensitivity factor is 1%. It is assumed that the pavement conditions on the highways are constant and the routing behaviours do not change, so the highway emissions do not respond to the eco-maintenance on the local road network.

As resources are diverted to roads with higher traffic rather than the worst conditions, the eco-maintenance strategy is not very effective in improving the average PCI of the local road network. After a 10-year analysis period, the average PCI under the eco-maintenance scenarios are found to be 3 to 7 points lower than the PCI-based maintenance, depending on the maintenance budget (3 points lower at low budget level and 7 points lower at the high budget level).

The effects of the eco-maintenance strategy are sensitive to the initial pavement conditions. When the starting average PCI of the local roads move from 79 (predicted by the degradation

model) to 74 (figures from the local agency), the cumulative savings due to eco-maintenance changes from 1.9% to 1.5% (high budget level, IRI sensitivity of 3%).

The effects of the eco-maintenance strategy depend on the maintenance gains. When the road works can recover only 75% or 50% of the loss of condition due to the pavement degradation, the 10-year cumulative emissions under the eco-routing scenario (high budget, IRI sensitivity of 3%) are 0.8% or 1.5% higher respectively compared to the 100% recovery case.

The effects of eco-maintenance are also sensitive to the pavement degradation rates. When the annual average pavement degradation rates are 1, 3 or 5 PCI, the 10-year cumulative savings due to the implementation of the eco-maintenance (high budget, IRI sensitivity of 3%) are 1.9%, 2.4% and 2.9%, compared to only implementing the PCI-based maintenance.

(b) The eco-routing strategy:

The effects of the eco-routing strategy in reducing the CO_2 emissions of the vehicular traffic depend highly on the eco-routing participation rates. On the network-level (including both the local roads and the highways), increasing eco-routing participation rates from 0% to 10%, 50% and 100% will result in a 0.7%, 3% and 5% savings in CO_2 emissions over a 10-year analysis period (high budget, IRI sensitivity of 3%).

Eco-routing is more effective in reducing the CO_2 emissions from the highway component than the local roads. Increasing the eco-routing participation rates from 0% to 10%, 50% and 100%, the CO_2 emission savings on the highways are 1.3%, 6.1% and 11.5% over a 10-year analysis period, while the savings are 0.4%, 1.7% and 0.9% for the local roads (high budget, IRI sensitivity of 3%).

The above results also depend on the budget level, as increasing budget on local road maintenance improves the pavement conditions, thus reducing the roughness-induced CO_2 emissions on the local roads. However, when the eco-routing participation rate is 100%, the improved local roads will also attract more vehicles to use them. The percentage of savings are less affected by the IRI sensitivity.

Variabilities in the CO_2 emission results given different random seeds (controlling which travellers are randomly selected to use eco-routing path) are very small.

Under the eco-routing scenario with participation rates of 50% and 100%, the total journey distances decrease. The 10-year average daily VKMT reduces by 2% and 5% compared to the shortest time routing case (all budget levels, all IRI sensitivity levels). While the reduction in VKMT with 10% eco-routing participation ratio is not obvious. However, the reduction in VKMT is not balanced in the whole network. On the highways, the 10-year average daily VKMT reduces by 1%, 5% and 12% from the highways compared to the shortest-time routing case with eco-routing participation ratios of 10%, 50% and 100%. On

the other side, the VKMT on the local roads increases by 0.5%, 1.8% and 3% compared to the shortest-time routing case with eco-routing participation ratios of 10%, 50% and 100%.

The increasing usage of the local roads also creates more congestions on this part of the network. The 10-year average daily VHT on the local roads increase by 0.4%, 4.7% and 21% compared to the shortest-time routing case with eco-routing participation ratios of 10%, 50% and 100%. Due to the use of the BPR volume-delay curve, the delays are penalised to the power of 4 with a unit increase in the traffic volumes, thus leading to the great increase in congestion when the vehicle concentration on local roads increase in the 100% eco-routing case.

The eco-routing strategy itself does not cause a change in the pavement degradation trend, due to the lack of interaction found between vehicle loading and the pavement degradations in the previous chapter.

(c) The combined effect of eco-routing and eco-maintenance:

The effect of eco-maintenance, when implemented alone on the local network, has comparable performance in reducing the network-wide CO_2 emissions as the eco-routing strategy with a participation rate of 10%. However, the savings in CO_2 emissions due to the eco-routing come mainly from the reduction of travel distance, especially on the highway parts of the network.

The combined emission savings of eco-maintenance and eco-routing when implemented simultaneously are the greatest among all scenarios. The combined savings in CO_2 emissions of the whole network are not significantly different from summing up savings of these two emission mitigation strategies implemented individually. However, when partitioning the savings to the highway component and the local roads component, it can be seen that the combined effects of eco-maintenance and eco-routing are not the simple sum of any individual approach, especially when the eco-routing participation ratio is high. The combined eco-routing (100% participation) and eco-maintenance will lead to a further 1.2% reduction in the 10-year average highway CO_2 emissions on top of the 11.5% from implementing eco-routing alone (high budget, IRI sensitivity of 3%), while eco-maintenance itself does not lead to savings in highway emissions. However, it is acknowledged that such reinforcing effects are quite weak for most of the simulation scenarios, e.g., when the eco-routing participation rate is low or a low IRI sensitivity factor is assumed.

Similarly, the 10-year average daily VHT and VKMT of the combined eco-maintenance and eco-routing are largely predicted by the values of eco-routing alone (eco-maintenance itself does not affect the traffic patterns), with the exceptions being the 100% eco-routing case. When the eco-routing participation ratio reaches 100%, the implementation of the eco-maintenance will lead to better pavement conditions on the local roads, thus attracting

travellers to this part of the network. For example, with the combined eco-maintenance and eco-routing, the average journey duration increases by 2.5% compared with just implementing eco-routing (high budget, IRI sensitivity of 3%, 100% eco-routing participation). Still, such changes in the traffic operation efficiency are rather small due to the interactions of eco-routing and eco-maintenance.

The final local road PCI conditions when implementing both eco-maintenance and eco-routing are similar to those obtained with just implementing eco-maintenance. As the PCI in this study is directly influenced only by the maintenance schedules, such results suggests that eco-maintenance, whether implemented individually or combined with eco-routing, leads to lower PCI compared to the PCI-based maintenance site selections.

(d) Traffic growth:

Traffic growth scenarios are analysed and presented toward the end of Chapter 6. Considering a gentle traffic growth ratio of around 1% per year for each district, the average daily trip counts are 4.6% higher during a 10-year analysis period compared to the no traffic growth case. Due to the increased number of trips, the traffic AAD-VHT and AAD-VKMT also increase as expected. Specifically, the network-level AAD-VKMT increases by about 4% and the AAD-VHT increases superlinearly by about 7-10%, due to the power law penalty on travel time given unit traffic volume increase in the BPR curve. The increase in the network-wide AAD-CO_2 can be as high as 5.4%, equivalent to the savings that are achieved with 100% eco-routing. But it should be highlighted that the traffic growth scenarios do not take into consideration the changes in the vehicle fleet compositions or fuel type changes, which may contribute positively to the eco-friendliness in the transportation system in the long term. In addition, as no traffic load induced pavement degradation is considered in this study, the traffic growth has almost no impact on the average PCI of the local roads, although the redistributions of the traffic do have an influence on the pavement maintenance scheduling of individual streets.

7.3 Policy implications

Various outcomes from this thesis may be informative to the policy making in transportation system management. To start with, high levels of eco-routing (with a participation rate of 50% or more) lead to significant ($\geq$ 3%) savings in terms of network-wide CO_2 emissions. However, this comes at a cost of the traffic efficiency. When the eco-routing participation rate is 50%, the average journey duration increases by 5%, which translates to about 2 minutes in absolute term. This may be considered acceptable, as an additional 2 minutes could be lost or gained at traffic intersections or the parking lots. However, when the eco-

routing participation rate reaches 100%, the average journey duration increases by 25%, or 10 minutes. Even though it leads to the greatest reduction in CO_2 emissions (5%), this compromises the economic aspects, such as the working hours. Theoretically, an eco-routing participation rate less than 50% is reasonable. While it is unclear how willingly the travellers will adopt this behaviour in reality. In this sense, additional studies on the real-life barriers or incentives related to eco-routing should also be conducted.

The benefit of eco-maintenance in terms of CO_2 emissions is comparable to that of 10% eco-routing. However, it should be noted that the adoption of eco-maintenance will inevitably direct the pavement repair funding to busy roads that carry higher volumes of traffic. While in real practice, other factors should also be considered, such as the presence of public transits or equitable distributions of resources.

Furthermore, the results of the combined eco-maintenance and eco-routing scenario show weak interactions between the two strategies except for the extreme case with 100% eco-routing. This implies that the two strategies can be studied separately for preliminary evaluations of the benefits of the carbon mitigation or environmental sustainability policies. However, for detailed assessments of policy or project-level outcomes, it is still recommended to consider both aspects together for more rigorous results.

In Chapter 4, it is found that the minimum probe ratio required to keep a satisfying traffic efficiency level is very low, at only 0.5% or 1%. It is acknowledged that the static traffic assignment (STA) procedure used for the traffic simulation has its problems in analysing highly dynamic traffic situations. However, the STA procedure is reasonably applicable when the traffic conditions do not change fast, such as during the off-peak hours. In order to balance the trade-offs between the network efficiency and user privacy, it might be useful to separate the system operations based on the traffic conditions. During off-peak hours, only probe data from public vehicles, e.g., taxis, should be sufficient to provide the required level of information.

Lastly, the study presented in this thesis would not be possible without the traffic and pavement condition input data being available. While on the other hand, the quality of some data and the missing information also pose constraints in the analyses. For example, it is difficult to extract a good degradation trend from the scattered pavement condition observations. The policy implication here is on the importance of the traffic and infrastructure condition monitoring, which may lead to more sophisticated or reliable outcomes of carbon mitigation analysis in the future.

7.4 Recommendations for further work

There are many elements of this thesis that can be further researched, whether within the same case study area or in terms of applying the methodology to other cities. Some directions for further improvements are identified as the following:

Compared with the existing studies in the literature, the traffic simulation model adopted in this thesis is still simplistic in many aspects. For example, activity-based travel demand generation can be adopted to replace the fixed trip-level travel demand of the TNC companies. The benefits of implementing an activity-based demand model include the greater flexibility in modelling individual decisions, such as electric vehicle own rates, working from home or avoiding to travel during the peak hours. These behaviours all have important implications in building a more eco-friendly transportation network.

Related to the activity-based demand generation approach is the multi-modal traffic network simulation. For transportation network-wide carbon emission analysis, it is also important to consider the role of the public transits.

Also, the authors would like to improve the temporal resolution of the traffic simulations by moving towards the dynamic traffic assignment (DTA) approach. Dynamic traffic simulations require checking the consistencies of the travel demand and network supply more frequently. Besides, the supply-demand interactions are also different from those adopted for the static assignment procedure. It is believed that the DTA approach, though more time consuming, would lead to more realistic sub-hourly traffic distribution outcomes.

Though the traffic simulation module is efficient for city-scale analysis, the performance should ideally be further improved so as to allow day-by-day maintenance disruption impact analysis. The inclusion of activity-based travel demand generation and the DTA would inevitably make the computation more expensive. However, as currently only 32 parallel threads are used for the traffic assignment calculation, the parallel computing ability of the HPC can still be further explored to distribute the work load to more parallel processes.

In terms of the pavement degradation analysis, it is recognised through the data analysis exercise in this study that the availability and quality of the infrastructure asset management data pose great limitations on the modelling options. Thus, one potential directions for future research is to investigate these impacts more systematically and study how the improvements in infrastructure sensing and asset management data collection may affect the understanding of the infrastructure performance.

The eco-maintenance scenario in this thesis assumes only one option of the maintenance technique, i.e., the road repaving. Besides, the material, transport and construction related emissions are not considered. To make the analysis of the eco-maintenance scenarios more

complete, other maintenance options, material choices and road work related factors can be included in future research.

References

[1] K. Ahn and H. A. Rakha. Network-wide impacts of eco-routing strategies: a large-scale case study. *Transportation Research Part D: Transport and Environment*, 25: 119–130, 2013. doi: https://doi.org/10.1016/j.trd.2013.09.006.

[2] K. Ahn, H. Rakha, A. Trani, and M. Van Aerde. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. *Journal of Transportation Engineering*, 128(2):182–190, 2002. doi: https://doi.org/10.1061/(ASCE)0733-947X(2002)128:2(182).

[3] Aimsun Software. About Aimsun, 2019. [online] https://www.aimsun.com. Accessed Jun 24, 2019.

[4] N. Alaswadko, R. Hassan, D. Meyer, and B. Mohammed. Modelling roughness progression of sealed granular pavements: a new approach. *International Journal of Pavement Engineering*, 20(2):222–232, 2019. doi: https://doi.org/10.1080/10298436.2017.1283689.

[5] S.-H. An, B.-H. Lee, and D.-R. Shin. A survey of intelligent transportation systems. In *2011 Third International Conference on Computational Intelligence, Communication Systems and Networks*, pages 332–337. IEEE, 2011. doi: https://doi.org/10.1109/CICSyN.2011.76.

[6] A. R. Andrade and P. F. Teixeira. Statistical modelling of railway track geometry degradation using Hierarchical Bayesian models. *Reliability Engineering & System Safety*, 142:169–183, 2015. doi: https://doi.org/10.1016/j.ress.2015.05.009.

[7] M. Anyala, J. Odoki, and C. Baker. Hierarchical asphalt pavement deterioration model for climate impact studies. *International Journal of Pavement Engineering*, 15(3): 251–266, 2014. doi: https://doi.org/10.1080/10298436.2012.687105.

[8] ARA, Inc. Guide for Mechanistic-Empirical design of new and rehabilitated pavement structures Part 3: Design analysis, 2004. [online] http://onlinepubs.trb.org/onlinepubs/archive/mepdg/guide.htm. Accessed Jun 24, 2019.

[9] V. T. Arasan and G. Dhivya. Measuring heterogeneous traffic density. In *Proceedings of International Conference on Sustainable Urban Transport and Environment*, Bangkok, 2008. World Academy of Science, Engineering and technology. [online] https://waset.org/publications/8440/measuring-heterogeneous-traffic-density. Accessed Jun 24, 2019.

[10] S. A. Arhin, L. N. Williams, A. Ribbiso, and M. F. Anderson. Predicting pavement condition index using international roughness index in a dense urban area. *Journal of Civil Engineering Research*, 5(1):10–17, 2015. doi: https://doi.org/10.5923/j.jce.20150501.02.

[11] A. Aw and M. Rascle. Resurrection of second order models of traffic flow. *SIAM Journal on Applied Mathematics*, 60(3):916–938, 2000. doi: https://doi.org/10.1137/S0036139997332099.

[12] J. E. Baerwald, M. J. Huber, and L. E. Keefer. *Transportation and traffic engineering handbook*. 1976. ISBN 978-1-118-76230-1.

[13] M. Balmer, K. Meister, K. Nagel, and K. Axhausen. Agent-based simulation of travel demand: Structure and computational performance of MATSim-T, 2008. [online] http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.542.2927&rep=rep1&type=pdf. Accessed Jun 24, 2019.

[14] M. Balmer, M. Rieser, K. Meister, D. Charypar, N. Lefebvre, and K. Nagel. MATSim-T: architecture and simulation times, 2009. [online] https://pdfs.semanticscholar.org/d7ec/e13bfd495ab4844d822a3b0d8ed5d5c5f926.pdf. Accessed Jun 24, 2019.

[15] M. Barth and K. Boriboonsomsin. Real-world carbon dioxide impacts of traffic congestion. *Transportation Research Record*, 2058(1):163–171, 2008. doi: https://doi.org/10.3141/2058-20.

[16] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz. Sumo - Simulation of Urban MObility: an overview, 2011. [online] citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.452.3969. Accessed Jun 24, 2019.

[17] M. Ben-Akiva, M. Bierlaire, H. Koutsopoulos, and R. Mishalani. DynaMIT: a simulation-based system for traffic prediction. In *DACCORD Short Term Forecasting Workshop*, pages 1–12. Delft The Netherlands, 1998. [online] http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.37.1549. Accessed Jun 24, 2019.

[18] J. Besag, J. York, and A. Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, 43(1):1–20, 1991. doi: https://doi.org/10.1007/BF00116466.

[19] S. Bessler and T. Paulin. Literature study on the state of the art of probe data systems in Europe, 2013. [online] http://www.fot-net.eu/download/fcd-report_final.pdf. Accessed Jun 24, 2019.

[20] C. Bester and W. Meyers. Saturation flow rates, 2007. [online] https://repository.up.ac.za/bitstream/handle/2263/5838/002.pdf?sequence=1&isAllowed=y. Accessed Jun 24, 2019.

[21] A. Black. The Chicago area transportation study: A case study of rational planning. *Journal of Planning Education and Research*, 10(1):27–37, 1990. doi: https://doi.org/10.1177/0739456X9001000105.

[22] M. Blangiardo and M. Cameletti. Bayesian computing. In *Spatial and spatio-temporal Bayesian models with R-INLA*, chapter 4, pages 75–125. John Wiley & Sons, Chichester, UK, 2015. doi: https://doi.org/10.1002/9781118950203.

[23] M. Blangiardo, M. Cameletti, G. Baio, and H. Rue. Spatial and spatio-temporal models with R-INLA. *Spatial and Spatio-temporal Epidemiology*, 4:33 – 49, 2013. ISSN 1877-5845. doi: https://doi.org/10.1016/j.sste.2012.12.001.

[24] M. C. Bliemer. Dynamic queuing and spillback in analytical multiclass dynamic network loading model. *Transportation Research Record*, 2029(1):14–21, 2007. doi: https://doi.org/10.3141/2029-02.

[25] M. C. J. Bliemer and H. Taale. Route generation and dynamic traffic assignment for large networks. In *Proceedings of the First International Symposium on Dynamic Traffic Assignment*, pages 90–99, 2006. [online] https://www.researchgate.net/publication/272497797_Route_Generation_and_Dynamic_Traffic_Assignment_for_Large_Networks. Accessed Jun 24, 2019.

[26] G. Boeing. R-tree spatial indexing with python, 2016. [online] https://geoffboeing.com/2016/10/r-tree-spatial-index-python/. Accessed Jun 24, 2019.

[27] G. Boeing. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65: 126–139, 2017. doi: https://doi.org/10.1016/j.compenvurbsys.2017.05.004.

[28] K. Boriboonsomsin and M. Barth. Impacts of road grade on fuel consumption and carbon dioxide emissions evidenced by use of advanced navigation systems. *Transportation Research Record: Journal of the Transportation Research Board*, (2139): 21–30, 2009. doi: https://doi.org/10.3141/2139-03.

[29] K. Boriboonsomsin, M. J. Barth, W. Zhu, and A. Vu. Eco-routing navigation system based on multisource historical and real-time traffic information. *IEEE Transactions on Intelligent Transportation Systems*, 13(4):1694–1704, 2012. doi: https://doi.org/10.1109/TITS.2012.2204051.

[30] J. L. Bowman and M. E. Ben-Akiva. Activity-based disaggregate travel demand model system with activity schedules. *Transportation research part a: policy and practice*, 35(1):1–28, 2001. doi: https://doi.org/10.1016/S0965-8564(99)00043-9.

[31] J. Brady and M. O'Mahony. Travel to work in Dublin. the potential impacts of electric vehicles on climate change and urban air quality. *Transportation Research Part D: Transport and Environment*, 16(2):188–193, 2011. doi: https://doi.org/10.1016/j.trd.2010.09.006.

[32] G. Brundtland, M. Khalid, S. Agnelli, S. Al-Athel, B. Chidzero, L. Fadika, V. Hauff, I. Lang, M. Shijun, M. M. de Botero, et al. Our common future (The Brundtland report). 1987.

[33] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis*, 28 (4):281–298, 1996. doi: https://doi.org/10.1111/j.1538-4632.1996.tb00936.x.

[34] Bureau of Public Rroads. Traffic assignment manual, 1964.

[35] C. Busch, M. Holst, and A. Christiansen. Identification and selection of pavement performance models, 2010. [online] https://www.researchgate.net/publication/282867666_NordFoU_PPM_-_Identification_and_Selection_of_Pavement_Performance_Models. Accessed Jun 24, 2019.

[36] California Air Resources Board. California greenhouse gas emissions from 2000 to 2016: Trends of emissions and other indicators, 2018. [online] https://www.arb.ca.gov/cc/inventory/pubs/reports/2000_2016/ghg_inventory_trends_00-16.pdf. Accessed Jan 15, 2019.

[37] CalTrans. State of the pavement report, 2015. [online] http://www.dot.ca.gov/hq/maint/Pavement/Offices/Planning_Programming/PDF/2015_SOP-7-9_12-22-15_FINAL_revised_1-4-15.docx. Accessed Jan 15, 2019.

[38] CalTrans. Pavement engineering considerations. In *Highway design manual*, chapter 610. Caltrans Publication Distribution Unit, Sacramento, USA, 2017. [online] http://www.dot.ca.gov/design/manuals/hdm/chp0610.pdf. Accessed Jan 15, 2019.

[39] J. L. Caradonna. *Sustainability: A history*. Oxford University Press, 2014. ISBN 978-1522663195.

[40] G. Casey. *Investigating the performance of transport infrastructure using real-time data and a scalable multi-modal agent based model*. Phd thesis, University of Cambridge, 2019.

[41] J. Castiglione, J. Freedman, and M. Bradley. Systematic investigation of variability due to random simulation error in an activity-based microsimulation forecasting model. *Transportation Research Record*, 1831(1):76–88, 2003. doi: https://doi.org/10.3141/1831-09.

[42] H. Castillo and D. E. Pitfield. ELASTIC - A methodological framework for identifying and selecting sustainable transport indicators. *Transportation Research Part D: Transport and Environment*, 15(4):179–188, 2010. doi: https://doi.org/10.1016/j.trd.2009.09.002.

[43] N. Cetin, A. Burri, and K. Nagel. A large-scale agent-based traffic microsimulation based on queue model, 2003. [online] http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.87.4169. Accessed Jun 24, 2019.

[44] D. Charypar, K. W. Axhausen, and K. Nagel. An event-driven parallel queue-based microsimulation for large scale traffic scenarios. *Arbeitsberichte Verkehrs-und Raumplanung*, 425, 2007. doi: https://doi.org/10.3929/ethz-a-005371338.

[45] K. Chatti and I. Zaabar. Estimating the effects of pavement condition on vehicle operating costs, 2012. doi: https://doi.org/10.17226/22808.

[46] M. V. Chester and A. Horvath. Environmental assessment of passenger transportation should include infrastructure and supply chains. *Environmental research letters*, 4(2):024008, 2009. doi: https://doi.org/10.1088/1748-9326/4/2/024008.

[47] C.-T. Chiu, T.-H. Hsu, and W.-F. Yang. Life cycle assessment on using recycled materials for rehabilitating asphalt pavements. *Resources, conservation and recycling*, 52(3):545–556, 2008. doi: https://doi.org/10.1016/j.resconrec.2007.07.001.

[48] City and County of San Francisco. Roadway repaving and street safety bond, 2011. [online] https://sfpublicworks.org/sites/default/files/1322-2011%20Road%20Repaving%20and%20Street%20Safety%20Bond%20Report%206-15.pdf. Accessed Jun 24, 2019.

[49] City and County of San Francisco. Roadway repaving and street safety bond, 2014. [online] https://cgoboc.sfgov.org/models/data/27Mar2014/docs/2011%20Roadway%20Repaving%20and%20Street%20Safety%20Bond.pdf. Accessed Jun 24, 2019.

[50] City and County of San Francisco. Current supervisor districts, 2019. [online] https://data.sfgov.org/Geographic-Locations-and-Boundaries/Current-Supervisor-Districts/8nkz-x4ny. Accessed Jun 24, 2019.

[51] S. Clark and D. Watling. Modelling network travel time reliability under stochastic demand. *Transportation Research Part B: Methodological*, 39(2):119–140, 2005. doi: https://doi.org/10.1016/j.trb.2003.10.006.

[52] S. Çolak, A. Lima, and M. C. González. Understanding congested travel in urban areas. *Nature communications*, 7:10793, 2016. doi: https://doi.org/10.1038/ncomms10793.

[53] A. Collop and D. Cebon. Modelling whole-life pavement performance. *Road Transport Technology*, 4:201–212, 1995. doi: https://doi.org/10.1243/PIME_PROC_1995_209_170_02.

[54] ConnectSF. About ConnectSF, 2019. [online] https://connectsf.org/about/about-connectsf/. Accessed Jun 24, 2019.

[55] ConnectSF. Daily trip-making patterns, 2019. [online] https://connectsf-trippatterns.sfcta.org. Accessed Jun 24, 2019.

[56] M. A. Costagliola, M. Costabile, and M. V. Prati. Impact of road grade on real driving emissions from two Euro 5 diesel vehicles. *Applied Energy*, 231:586–593, 2018. doi: https://doi.org/10.1016/j.apenergy.2018.09.108.

[57] S. B. Costello, M. S. Snaith, H. Kerali, L. V. Tachtsi, and J. J. Ortiz-García. Stochastic model for strategic assessment of road maintenance. In *Proceedings of the Institution of Civil Engineers-Transport*, volume 158, pages 203–211. Thomas Telford Ltd, 2005. doi: https://doi.org/10.1680/tran.2005.158.4.203.

[58] C. Coupe, H. Le Maitre, and S. Benelli. Empirical estimation of the variability of travel time. *Transportation research procedia*, 25:2769–2783, 2017. doi: https://doi.org/10.1016/j.trpro.2017.05.225.

[59] Y. Croissant, G. Millo, et al. Panel data econometrics in R: The plm package. *Journal of Statistical Software*, 27(2):1–43, 2008. doi: https://doi.org/10.18637/jss.v027.i02.

[60] S. Dafermos. Traffic equilibrium and variational inequalities. *Transportation science*, 14(1):42–54, 1980. doi: https://doi.org/10.1287/trsc.14.1.42.

[61] C. F. Daganzo. Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B: Methodological*, 29(4):277–286, 1995. doi: https://doi.org/10.1016/0191-2615(95)00007-Z.

[62] W. Davidson, R. Donnelly, P. Vovsha, J. Freedman, S. Ruegg, J. Hicks, J. Castiglione, and R. Picado. Synthesis of first practices and operational research approaches in activity-based travel demand modeling. *Transportation Research Part A: Policy and Practice*, 41(5):464–488, 2007. doi: https://doi.org/10.1016/j.tra.2006.09.003.

[63] C. de Fabritiis, R. Ragona, and G. Valenti. Traffic estimation and prediction based on real time floating car data. In *2008 11th International IEEE Conference on Intelligent Transportation Systems*, pages 197–203, Oct 2008. doi: https://doi.org/10.1109/ITSC.2008.4732534.

[64] S. Debnath, N. Ganguly, and P. Mitra. Feature weighting in content based recommendation system using social network analysis. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, pages 1041–1042, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-085-2. doi: 10.1145/1367497.1367646.

[65] M. M. Deshmukh. Development of equations to determine the increase in pavement condition due to treatment and the rate of decrease in condition after treatment for a local agency pavement network. Master's thesis, Texas A & M University, 2010. [online] http://hdl.handle.net/1969.1/ETD-TAMU-2009-05-317. Accessed Jun 24, 2019.

[66] S. A. Dewan and R. E. Smith. Estimating international roughness index from pavement distresses to calculate vehicle operating costs for the SF bay area. *Transportation research record*, 1816(1):65–72, 2002. doi: https://doi.org/10.3141/1816-08.

[67] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959. doi: https://doi.org/10.1007/BF01386390.

[68] C. F. Dormann, J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, D. W. Kissling, I. Kühn, R. Ohlemüller, P. R. Peres-Neto, B. Reineking, B. Schröder, F. M. Schurr, and R. Wilson. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, 30 (5):609–628, 2007. ISSN 1600-0587. doi: https://doi.org/10.1111/j.2007.0906-7590.05171.x.

[69] R. Dowling, P. Ryus, B. Schroeder, M. Kyte, F. T. Creasey, N. Rouphail, A. Hajbabaie, and D. Rhoades. Planning and Preliminary Engineering Applications Guide to the Highway Capacity Manual, 2016. [online] http://www.trb.org/PlanningForecasting/Blurbs/174958.aspx. Accessed Jun 24, 2019.

[70] EIA. Fuel economy improvements are projected to reduce future gasoline use, 2017. [online] https://www.eia.gov/todayinenergy/detail.php?id=31332. Accessed Jun 24, 2019.

[71] EPA. Sources of greenhouse gas emissions: Transportation, 2018. [online] https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions. Accessed Jun 24, 2019.

[72] T. Ercan, M. Kucukvar, O. Tatari, and H. Al-Deek. Congestion relief based on intelligent transportation systems in Florida: analysis of triple bottom line sustainability impact. *Transportation Research Record*, 2380(1):81–89, 2013. doi: https://doi.org/10.3141/2380-09.

[73] G. Erhardt, P. Brinckerhoff, D. Ory, A. Sarvepalli, J. Freedman, J. Hood, and B. Stabler. MTC's Travel Model One: Applications of an activity-based model in its first year. In *5th Transportation Research Board Innovations in Travel Modeling Conference*, 2012. [online] https://www.researchgate.net/publication/262261101_MTC's_Travel_Model_One_Applications_of_an_Activity-Based_Model_in_its_First_Year. Accessed Jun 24, 2019.

[74] European Commission. PARIS: Performance Analysis of Road InfraStructure, 1999.

[75] European Commmitte for Standardization, Technical Committe 350. *BS EN 15643-5:2017 - Framework on specific principles and requirement for civil engineering works*. 2017. ISBN 9780580925009.

[76] S. Fan. Generic second order models, 2019. [online] http://publish.illinois.edu/shimao-fan/research/generic-second-order-models/. Accessed Jun 24, 2019.

[77] E. Ferkingstad and H. Rue. Improving the INLA approach for approximate Bayesian inference for latent Gaussian models. *Electronic Journal of Statistics*, 9(2):2706–2731, 2015. doi: https://doi.org/10.1214/15-EJS1092.

[78] P. Fernandes and U. Nunes. Platooning of autonomous vehicles with intervehicle communications in sumo traffic simulator. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 1313–1318. IEEE, 2010. doi: https://doi.org/10.1109/ITSC.2010.5625277.

[79] A. Ferreira, L. Picado-Santos, and A. Antunes. A segment-linked optimization model for deterministic pavement management systems. *International Journal of Pavement Engineering*, 3(2):95–105, 2002. doi: https://doi.org/10.1080/10298430290030603.

[80] A. Ferreira, L. de Picado-Santos, Z. Wu, and G. Flintsch. Selection of pavement performance models for use in the Portuguese PMS. *International Journal of Pavement Engineering*, 12(1):87–97, 2011. doi: https://doi.org/10.1080/10298436.2010.506538.

[81] M. Ferreira and P. M. d'Orey. On the impact of virtual traffic lights on carbon emissions mitigation. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):284–295, 2012. doi: https://doi.org/10.1109/TITS.2011.2169791.

[82] Food and Agriculture Organization of the United States. Crop rotation, 2019. [online] http://www.fao.org/tc/exact/sustainable-agriculture-platform-pilot-website/nutrients-and-soil-fertility-management/crop-rotation/en/. Accessed Jun 24, 2019.

[83] F. Galatioto, Y. Huang, T. Parry, R. Bird, and M. Bell. Traffic modelling in system boundary expansion of road pavement life cycle assessment. *Transportation Research Part D: Transport and Environment*, 36:65–75, 2015. doi: https://doi.org/10.1016/j.trd.2015.02.007.

[84] E. Galli, L. Cuéllar, S. Eidenbenz, M. Ewers, S. Mniszewski, and C. Teuscher. Activitysim: large-scale agent-based activity generation for infrastructure simulation. In *Proceedings of the 2009 spring simulation multiconference*, page 16. Society for Computer Simulation International, 2009. [online] https://dl.acm.org/citation.cfm?id=1639826&dl=ACM&coll=DL. Accessed Jun 24, 2019.

[85] J. Gallus, U. Kirchner, R. Vogt, and T. Benter. Impact of driving style and road grade on gaseous exhaust emissions of passenger vehicles measured by a portable emission measurement system (PEMS). *Transportation Research Part D: Transport and Environment*, 52:215–226, 2017. doi: https://doi.org/10.1016/j.trd.2017.03.011.

[86] L. Gao and Z. Zhang. Approximation approach to problem of large-scale pavement maintenance and rehabilitation. *Transportation Research Record: Journal of the Transportation Research Board*, 2304:112–118, 2012. doi: https://doi.org/10.3141/2304-13.

[87] V. V. Gayah and V. V. Dixit. Using mobile probe data and the macroscopic fundamental diagram to estimate network densities: Tests using microsimulation. *Transportation Research Record*, 2390(1):76–86, 2013. doi: https://doi.org/10.3141/2390-09.

[88] F. W. Geels. A socio-technical analysis of low-carbon transitions: introducing the multi-level perspective into transport studies. *Journal of Transport Geography*, 24: 471–482, 2012. doi: https://doi.org/10.1016/j.jtrangeo.2012.01.021.

[89] R. Geisberger, P. Sanders, D. Schultes, and D. Delling. Contraction hierarchies: Faster and simpler hierarchical routing in road networks. In C. C. McGeoch, editor, *Experimental Algorithms*, pages 319–333, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. ISBN 978-3-540-68552-4.

[90] A. Gelman. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, 1(3):515–534, 2006. doi: https://doi.org/10.1214/06-BA117A.

[91] A. Gelman and J. Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical Methods for Social Research. Cambridge University Press, 2006. doi: https://doi.org/10.1017/CBO9780511790942.

[92] P. G. Gipps. A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological*, 20(5):403–414, 1986. doi: https://doi.org/10.1016/0191-2615(86)90012-3.

[93] F. Giustozzi, M. Crispino, and G. Flintsch. Multi-attribute life cycle assessment of preventive maintenance treatments on road pavements for achieving environmental sustainability. *The International Journal of Life Cycle Assessment*, 17(4):409–419, 2012. doi: https://doi.org/10.1007/s11367-011-0375-6.

[94] K. Golabi, R. B. Kulkarni, and G. B. Way. A statewide pavement management system. *Interfaces*, 12(6):5–21, 1982. [online] https://www.jstor.org/stable/25060344. Accessed Jun 24, 2019.

[95] Google. The bright side of sitting in traffic: Crowdsourcing road congestion data, 2009. [online] https://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html. Accessed Feb 1, 2019.

[96] Google Maps. Google maps showing the real-time traffic in downtown SF, 2019. [online] https://goo.gl/maps/2pTx44KCx7MHbXsu6. Accessed Feb 1, 2019.

[97] D. L. Greene and S. E. Plotkin. Reducing greenhouse gas emissions from US transportation, 2011. [online] https://rosap.ntl.bts.gov/view/dot/23588. Accessed Jun 24, 2019.

[98] B. Griesenbeck and G. Garry. Comparison of activity-based tour model to four-step model as a tool for metropolitan transportation planning. In *Proceedings of the National Transportation Planning Applications Conference, May*, pages 6–10, 2007. [online] https://www.trbappcon.org/2007conf/papers/session08/05-Comparison_of_ActivityBased_Tour_Model_to_FourStep_Model.pdf. Accessed Jun 24, 2019.

[99] A. Guhnemann, R.-P. Schafer, K.-U. Thiessenhusen, and P. Wagner. Monitoring traffic and emissions by floating car data, 2004. [online] http://hdl.handle.net/2123/19267. Accessed Jun 24, 2019.

[100] S. W. Haider, G. Y. Baladi, K. Chatti, and C. M. Dean. Effect of frequency of pavement condition data collection on performance prediction. *Transportation Research Record*, 2153(1):67–80, 2010.

[101] F. L. Hall. Traffic stream characteristics, 1996. [online] https://www.researchgate.net/publication/242405119_Traffic_Stream_Characteristics. Accessed Jun 24, 2019.

[102] M. Hatzopoulou and E. J. Miller. Linking an activity-based travel demand model with traffic emission and dispersion models: Transport's contribution to air pollution in Toronto. *Transportation Research Part D: Transport and Environment*, 15(6):315–325, 2010. [online] https://doi.org/10.1016/j.trd.2010.03.007. Accessed Jun 24, 2019.

[103] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies*, 18(4):568–583, 2010. doi: https://doi.org/10.1016/j.trc.2009.10.006.

[104] HMEP. Life cycle planning toolkit user guidance: Deterioration models, 2012. [online] http://www.highwaysefficiency.org.uk/efficiency-resources/asset-management/life-cycling-planning-toolkit.html. Accessed Jun 24, 2019.

[105] P. R. Hobbs, K. Sayre, and R. Gupta. The role of conservation agriculture in sustainable agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1491):543–555, 2007. doi: https://doi.org/10.1098/rstb.2007.2169.

[106] A. Horvath and C. Hendrickson. Comparison of environmental implications of asphalt and steel-reinforced concrete pavements. *Transportation Research Record: Journal of the Transportation Research Board*, (1626):105–113, 1998. doi: https://doi.org/10.3141/1626-13.

[107] D. Howey, R. North, and R. Martines-Botas. Road transport technology and climate change mitigation, 2010. [online] https://www.imperial.ac.uk/media/imperial-college/grantham-institute/public/publications/briefing-papers/Road-transport-technology-and-climate-mitigation--Grantham-BP-2.pdf. Accessed Jun 24, 2019.

[108] Y. Huang, R. Bird, and M. Bell. A comparative study of the emissions by road maintenance works and the disrupted traffic using life cycle assessment and micro-simulation. *Transportation Research Part D: Transport and Environment*, 14(3): 197–204, 2009. doi: https://doi.org/10.1016/j.trd.2008.12.003.

[109] Y. H. Huang. *Pavement analysis and design*. Prentice-Hall, 1993. ISBN 9780136552758.

[110] J. D. Hunt and K. Stefan. Tour-based microsimulation of urban commercial movements. *Transportation Research Part B: Methodological*, 41(9):981–1013, 2007. doi: https://doi.org/10.1016/j.trb.2007.04.009.

[111] INLA documentation. Truncated Gaussian prior, 2010. [online] https://inla.r-inla-download.org/r-inla.org/doc/prior/prior-logtnorm.pdf. Accessed Jun 24, 2019.

[112] Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2014: Mitigation of Climate Change*. Cambridge University Press, 2014. [online] https://www.ipcc.ch/report/ar5/wg3/. Accessed Jun 24, 2019.

[113] International Energy Agency. CO2 emissions from fuel combustion 2018 highlights, 2018. [online] https://webstore.iea.org/co2-emissions-from-fuel-combustion-2018-highlights. Accessed Jun 24, 2019.

[114] S. Inti, S. A. Martin, and V. Tandon. Necessity of including maintenance traffic delay emissions in life cycle assessment of pavements. *Procedia Engineering*, 145:972–979, 2016. doi: https://doi.org/10.1016/j.proeng.2016.04.126.

[115] S. Islam and W. G. Buttlar. Effect of pavement roughness on user costs. *Transportation Research Record*, 2285(1):47–55, 2012. doi: https://doi.org/10.3141/2285-06.

[116] S. Jahanbakhsh, L. Gao, and Z. Zhang. Estimating spatial dependence associated with deterioration process of road network. Technical Report 16-6179, Transport Research Board, Washington DC, USA, 2016. [online] https://trid.trb.org/view/1394311. Accessed Jun 24, 2019.

[117] M. Jakimavičius and M. Burinskiene. Assessment of vilnius city development scenarios based on transport system modelling and multicriteria analysis. *Journal of Civil Engineering and management*, 15(4):361–368, 2009. doi: DOI:10.3846/1392-3730.2009.15.361-368.

[118] R. Jayakrishnan, H. S. Mahmassani, and T.-Y. Hu. An evaluation tool for advanced traffic information and management systems in urban networks. *Transportation Research Part C: Emerging Technologies*, 2(3):129–147, 1994. doi: https://doi.org/10.1016/0968-090X(94)90005-1.

[119] C. M. Jeon and A. Amekudzi. Addressing sustainability in transportation systems: definitions, indicators, and metrics. *Journal of Infrastructure Systems*, 11(1):31–50, 2005. doi: https://doi.org/10.1061/(ASCE)1076-0342(2005)11:1(31).

[120] T. Jeske. Floating car data from smartphones: What Google and Waze know about you and how hackers can control traffic, 2013. [online] https://media.blackhat.com/eu-13/briefings/Jeske/bh-eu-13-floating-car-data-jeske-slides.pdf. Accessed Jun 24, 2019.

[121] K. D. Johnson and K. A. Cation. Performance prediction development using three indexes for North Dakota pavement management system. *Transportation Research Record*, 1344, 1992. [online] http://onlinepubs.trb.org/Onlinepubs/trr/1992/1344/1344-004.pdf. Accessed Jun 24, 2019.

[122] J. W. Joubert, P. J. Fourie, and K. W. Axhausen. Large-scale agent-based combined traffic simulation of private cars and commercial vehicles. *Transportation Research Record*, 2168(1):24–32, 2010. doi: https://doi.org/10.3141/2168-04.

[123] N. Kai. Is MATSim microscopic or mesoscopic?, 2018. [online] https://github.com/matsim-org/matsim-code-examples/wiki/faq-114524171. Accessed Jun 24, 2019.

[124] H. G. Kerali. The highway development and management series vol. 1: Overview of HDM-4, 2000. [online] http://piarc.rmto.ir/DocLib4/Volume1.pdf. Accessed Jun 24, 2016.

[125] B. Kerner, C. Demir, R. Herrtwich, S. Klenov, H. Rehborn, M. Aleksic, and A. Haug. Traffic state detection with floating car data in road networks. In *Proceedings of IEEE Intelligent Transportation Systems*, pages 44–49. IEEE, 2005. doi: https://doi.org/10.1109/ITSC.2005.1520133.

[126] A. Kesting, M. Treiber, and D. Helbing. General lane-changing model MOBIL for car-following models. *Transportation Research Record*, 1999(1):86–94, 2007. doi: https://doi.org/10.3141/1999-10.

[127] J. Kim and H. S. Mahmassani. A finite mixture model of vehicle-to-vehicle and day-to-day variability of traffic network travel times. *Transportation Research Part C: Emerging Technologies*, 46:83–97, 2014. doi: https://doi.org/10.1016/j.trc.2014.05.011.

[128] J. Kim and H. S. Mahmassani. Compound Gamma representation for modeling travel time variability in a traffic network. *Transportation Research Part B: Methodological*, 80:40–63, 2015. doi: https://doi.org/10.1016/j.trb.2015.06.011.

[129] K. M. Kockelman and J. Ma. Freeway speeds and speed variations preceding crashes, within and across lanes. *Journal of the Transportation Research Forum*, 46(1):43–61, 2007. doi: http://dx.doi.org/10.5399/osu/jtrf.46.1.976.

[130] A. Kotsialos, M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham. Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET. *IEEE Transactions on Intelligent Transportation Systems*, 3(4):282–292, 2002. doi: http://dx.doi.org/10.1109/TITS.2002.806804.

[131] K. J. Krizek. Neighborhood services, trip purpose, and tour-based travel. *Transportation*, 30(4):387–410, 2003. doi: https://doi.org/10.1023/A:1024768007730.

[132] K. Kumar. Priority-queue based shortest path library, 2018. [online] https://github.com/cb-cities/sp. Accessed Jun 24, 2019.

[133] E. Kwon and S. Pitt. Evaluation of emergency evacuation strategies for downtown event traffic using a dynamic network model. *Transportation Research Record*, 1922 (1):149–155, 2005. doi: https://doi.org/10.1177/0361198105192200119.

[134] K. Lautso, K. Spiekermann, and M. Wegener. Modelling policies for urban sustainability, 2002. [online] http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa02/cd-rom/papers/384.pdf. Accessed Jun 24, 2019.

[135] M. L. Lavine and J. S. Hodges. On rigorous specification of iCAR models. *The American Statistician*, 66(1):42–49, 2012. doi: https://doi.org/10.1080/00031305.2012.654746.

[136] D. Levinson. The value of advanced traveler information systems for route choice. *Transportation Research Part C: Emerging Technologies*, 11(1):75–87, 2003. doi: https://doi.org/10.1016/S0968-090X(02)00023-2.

[137] J. W. Lichstein, T. R. Simons, S. A. Shriner, and K. E. Franzreb. Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, 72(3):445–463, 2002. ISSN 1557-7015. doi: https://doi.org/10.1890/0012-9615(2002)072[0445:SAAAMI]2.0.CO;2.

[138] J. Lidicker, N. Sathaye, S. Madanat, and A. Horvath. Pavement resurfacing policy for minimization of life-cycle costs and greenhouse gas emissions. *Journal of Infrastructure Systems*, 19(2):129–137, 2012. doi: https://doi.org/10.1061/(ASCE)IS.1943-555X.0000114.

[139] M. J. Lighthill and G. B. Whitham. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 229(1178):317–345, 1955. doi: https://doi.org/10.1098/rspa.1955.0089.

[140] M. Lin and W.-J. Hsu. Mining GPS data for mobility patterns: A survey. *Pervasive and Mobile Computing*, 12:1–16, 2014. doi: https://doi.org/10.1016/j.pmcj.2013.06.005.

[141] T. Litman. Developing indicators for comprehensive and sustainable transport planning. *Transportation Research Record*, 2017(1):10–15, 2007.

[142] T. Litman and D. Burwell. Issues in sustainable transportation. *International Journal of Global Environmental Issues*, 6(4):331–347, 2006. doi: https://doi.org/10.1504/IJGENVI.2006.010889.

[143] S. Lorkowski, P. Mieth, and R. Schäfer. New ITS applications for metropolitan areas based on floating car data. In *ECTRI Young Researcher Seminar, Den Haag*, 2005. [online] https://elib.dlr.de/21000/. Accessed Jun 24, 2019.

[144] A. Louhghalam, M. Akbarian, and F.-J. Ulm. Carbon management of infrastructure performance: Integrated big data analytics and pavement-vehicle-interactions. *Journal of Cleaner Production*, 142:956–964, 2017. doi: https://doi.org/10.1016/j.jclepro.2016.06.198.

[145] R. Ma, X. J. Ban, and W. Szeto. Emission modeling and pricing in dynamic traffic networks. *Transportation Research Procedia*, 9:106–129, 2015. doi: https://doi.org/10.1016/j.trpro.2015.07.007.

[146] S. Maerivoet and B. De Moor. Transportation planning and traffic flow models, 2005. [online] https://arxiv.org/abs/physics/0507127. Accessed Jun 24, 2019.

[147] H. S. Mahmassani and K. F. Abdelghany. Dynasmart-ip: Dynamic traffic assignment meso-simulator for intermodal networks. In W. Lam and M. Bell, editors, *Advanced Modeling for Transit Operations and Service Planning*, pages 200–229. Elsevier, 2002. doi: https://doi.org/10.1108/9780585475226-008.

[148] H. S. Mahmassani, T. Hou, and J. Dong. Characterizing travel time variability in vehicular traffic networks: deriving a robust relation for reliability analysis. *Transportation Research Record*, 2315(1):141–152, 2012. doi: https://doi.org/10.3141/2315-15.

[149] K. J. Malakorn and B. Park. Assessment of mobility, energy, and environment impacts of IntelliDrive-based cooperative adaptive cruise control and intelligent traffic signal control. In *Proceedings of the 2010 IEEE International Symposium on Sustainable Systems and Technology*, pages 1–6. IEEE, 2010. doi: https://doi.org/10.1109/ISSST.2010.5507709.

[150] M. G. McNally. *The Four-Step Model*. Handbook of Transportation Modelling. Emerald Group Publishing Limited, Bingley, UK, 2007. ISBN 978-0-08-045376-7.

[151] M. G. McNally and C. Rindt. The activity-based approach, 2008. [online] https://cloudfront.escholarship.org/dist/prd/content/qt86h7f5v0/qt86h7f5v0.pdf?t=lnrbjv. Accessed Jun 24, 2019.

[152] J. McQueen and D. Timm. Statistical analysis of automated versus manual pavement condition surveys. *Transportation Research Record*, 1940:53–62, 2005. doi: https://doi.org/10.1177/0361198105194000107.

[153] D. H. Meadows, D. H. Meadows, J. Randers, and W. W. Behrens III. The limits to growth: A report to the Club of Rome (1972), 1972. [online] https://www.clubofrome.org/report/the-limits-to-growth/. Accessed Jun 24, 2019.

[154] M. Miller. *Seismic risk assessment of complex transportation networks*. Phd thesis, Stanford University, 2014.

[155] E. Minge, J. Kotzenmacher, and S. Peterson. Evaluation of non-intrusive technologies for traffic detection, 2010. [online] https://www.pooledfund.org/documents/TPF-5_171/2010-36.pdf, Accessed Jun 24, 2019.

[156] Mobile Millennium Project. Mobile Millennium: History of the project, 2010. [online] https://traffic.berkeley.edu/project. Accessed Jun 24, 2019.

[157] A. Montanaro. Priority-queues and Dijkstra's algorithm, 2013. [online] https://people.maths.bris.ac.uk/~csxam/teaching/dsa/dijkstra.pdf. Accessed Jun 24, 2019.

[158] G. Morosiuk, M. Riley, and J. Odoki. Modelling road deterioration and works effects in HDM-4. Technical report, World Road Association (PIARC) and The World Bank, Paris, France and Washington DC, US, 2004.

[159] U.-M. Mroueh, P. Eskola, J. Laine-Ylijoki, K. Wellman, E. Mäkelä, M. Juvankoski, and A. Ruotoistenmäki. *Life cycle assessment of road construction*. Finnish National Road Administration, 1999. ISBN 951-726-633-2.

[160] MTC. Travel model development: Calibration and validation, 2012. [online] https://mtcdrive.app.box.com/s/7crr7bwhromi2au42jnpp11fqe5l24xq. Accessed Jun 24, 2019.

[161] MTC. Travel model data: Data repository, 2018. [online] http://data.mtc.ca.gov/data-repository/. Accessed Jun 24, 2019.

[162] K. Nagel, J. ESSER, and M. Rickert. Large-scale traffic simulations for transportation planning. In D. Stauffer, editor, *Annual Reviews of Computational Physics VII*, chapter 6, pages 151–202. World Scientific, Singapore, 2000. doi: https://doi.org/10.1142/4229.

[163] National Programme on Technology Enhanced Learning (NPTEL). Fundamental parameters of traffic flow, 2007. [online] https://nptel.ac.in/courses/105101087/downloads/Lec-30.pdf. Accessed Feb 2, 2019.

[164] National Programme on Technology Enhanced Learning (NPTEL). Traffic assignment, 2007. [online] https://nptel.ac.in/courses/105101087/downloads/Lec-10.pdf. Accessed Feb 2, 2019.

[165] National Research Council. *Automotive Fuel Economy: How Far Can We Go?* The National Academies Press, Washington, DC, 1992. ISBN 978-0-309-04530-8. doi: https://doi.org/10.17226/1806.

[166] National Research Council and others. Sustainability and the US EPA, 2011. [online] https://www.epa.gov/sustainability/sustainability-and-us-epa. Accessed Jun 24, 2019.

[167] G. F. Newell. A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B: Methodological*, 27(4):281–287, 1993. doi: https://doi.org/10.1016/0191-2615(93)90038-C.

[168] R. B. Noland and J. W. Polak. Travel time variability: A review of theoretical and empirical issues. *Transport Reviews*, 22(1):39–54, 2002. doi: https://doi.org/10.1080/01441640010022456.

[169] M. M. Nunez and M. Y. Shahin. Pavement condition data analysis and modeling. *Transportation Research Record*, 1070:125–132, 1986. [online] http://onlinepubs.trb.org/onlinepubs/trr/1986/1070/1070-015.pdf. Accessed Jun 24, 2019.

[170] J. B. Odoki, M. Anyala, and E. Bunting. HDM-4 adaptation for strategic analysis of UK local roads. In *Proceedings of the Institution of Civil Engineers-Transport*, volume 166, pages 65–78. Thomas Telford Ltd, 2013. doi: https://doi.org/10.1680/tran.9.00026.

[171] Open Data Commons. ODC Public Domain Dedication and Licence (PDDL), 2009. [online] https://opendatacommons.org/licenses/pddl/index.html. Accessed Jun 24, 2019.

[172] OpenStreetMap Wiki. About OpenStreetMap, 2017. URL https://wiki.openstreetmap.org/wiki/About_OpenStreetMap. [online] https://wiki.openstreetmap.org/wiki/About_OpenStreetMap.

[173] OpenStreetMap Wiki. Map features, 2018. [online] https://wiki.openstreetmap.org/wiki/Map_Features#Highway. Accessed Jun 24, 2019.

[174] D. Ory. Introducing Doppelganger, Model Lab's open-source population synthesizer, 2017. [online] https://www.sidewalklabs.com/blog/a-first-step-toward-creating-a-digital-planning-laboratory-is-populating-it/. Accessed Jun 24, 2019.

[175] Overpass Wiki. Overpass API, 2018. [online] https://wiki.openstreetmap.org/wiki/Overpass_API. Accessed Jun 24, 2019.

[176] M. Papageorgiou. Dynamic modeling, assignment, and route guidance in traffic networks. *Transportation Research Part B: Methodological*, 24(6):471–495, 1990. doi: https://doi.org/10.1016/0191-2615(90)90041-V.

[177] B.-J. Park, T. Kim, I. Yang, J. Heo, and B. Son. A method for measuring accurate traffic density by aerial photography. *Journal of Advanced Transportation*, 49(4): 568–580, 2015. doi: https://doi.org/10.1002/atr.1288.

[178] K. Park, Y. Hwang, S. Seo, and H. Seo. Quantitative assessment of environmental impacts on life cycle of highways. *Journal of Construction Engineering and Management*, 129(1):25–31, 2003. doi: https://doi.org/10.1061/(ASCE)0733-9364(2003)129:1(25).

[179] K. Park, N. E. Thomas, and K. Wayne Lee. Applicability of the international roughness index as a predictor of asphalt pavement condition. *Journal of Transportation Engineering*, 133(12):706–709, 2007. doi: https://doi.org/10.1061/(ASCE)0733-947X(2007)133:12(706).

[180] W. D. Paterson. *Road deterioration and maintenance effects: Models for planning and management*. The Word Bank, Washington DC, USA, 1987. ISBN 0801835909. [online] http://documents.worldbank.org/curated/en/222951468765265396/Road-deterioration-and-maintenance-effects-models-for-planning-and-management. Accessed Jun 24, 2019.

[181] Pavement Interactive. 1993 AASHTO flexible pavement structural design, 2008. [online] https://www.pavementinteractive.org/reference-desk/design/structural-design/1993-aashto-flexible-pavement-structural-design/. Accessed Jun 24, 2019.

[182] Pavement Interactive. Structural number, 2009. [online] http://www.pavementinteractive.org/article/structural-number/. Accessed Jun 24, 2019.

[183] Pavement Interactive. What is mechanistic-empirical design? – The MEPDG and you, 2012. [online] https://www.pavementinteractive.org/what-is-mechanistic-empirical-design-the-mepdg-and-you/. Accessed Jun 24, 2019.

[184] Pavement Interactive. Equivalent single axle load, 2018. [online] https://www.pavementinteractive.org/reference-desk/design/design-parameters/equivalent-single-axle-load/. Accessed Jun 24, 2019.

[185] H. J. Payne. FREFLO: A macroscopic simulation model of freeway traffic. *Transportation Research Record*, (722), 1979. [online] http://onlinepubs.trb.org/Onlinepubs/trr/1979/722/722-010.pdf. Accessed Jun 24, 2019.

[186] Pew Research Center. US smartphone use in 2015, 2015. [online] http://www.pewresearch.org/wp-content/uploads/sites/9/2015/03/PI_Smartphones_0401151.pdf. Accessed Jun 24, 2015.

[187] PTV Group. PTV Vissim, 2019. [online] http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/. Accessed Jun 24, 2019.

[188] M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014. ISBN 978-0471727828.

[189] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. [online] https://www.R-project.org/. Accessed Jun 24, 2019.

[190] M. Rahman, M. Chowdhury, Y. Xie, and Y. He. Review of microscopic lane-changing models and future research opportunities. *IEEE transactions on intelligent transportation systems*, 14(4):1942–1956, 2013. doi: 10.1109/TITS.2013.2272074.

[191] H. Rakha, M. Van Aerde, K. Ahn, and A. Trani. Requirements for evaluating traffic signal control impacts on energy and emissions based on instantaneous speed and acceleration measurements. *Transportation Research Record*, 1738(1):56–67, 2000. doi: https://doi.org/10.3141/1738-07.

[192] R. A. Ramirez-Flores and C. Chang-Albitres. A stochastic approach for pavement condition projections and budget needs for the MTC pavement management system, 2012. [online] http://onlinepubs.trb.org/onlinepubs/conferences/2012/assetmgmt/presentations/Data-A-Ramirez-Flores-Chang-Albitres.pdf. Accessed Jun 24, 2019.

[193] D. Reger, S. Madanat, and A. Horvath. The effect of agency budgets on minimizing greenhouse gas emissions from road rehabilitation policies. *Environmental Research Letters*, 10(11):114007, 2015. doi: https://doi.org/10.1088/1748-9326/10/11/114007.

[194] M. Reisi, L. Aye, A. Rajabifard, and T. Ngo. Transport sustainability index: Melbourne case study. *Ecological Indicators*, 43:288–296, 2014. doi: https://doi.org/10.1016/j.ecolind.2014.03.004.

[195] P. I. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956. doi: https://doi.org/10.1287/opre.4.1.42.

[196] W. Riggs. Dealing with parking issues on an urban campus: The case of UC Berkeley. *Case Studies on Transport Policy*, 2(3):168–176, 2014. doi: https://doi.org/10.1016/j.cstp.2014.07.009.

[197] T. F. Rossi and Y. Shiftan. Tour based travel demand modeling in the US. *IFAC Proceedings Volumes*, 30(8):381–386, 1997. doi: https://doi.org/10.1016/S1474-6670(17)43853-5.

[198] H. Rue and S. Martino. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. *Journal of Statistical Planning and Inference*, 137(10):3177 – 3192, 2007. ISSN 0378-3758. doi: https://doi.org/10.1016/j.jspi.2006.07.016. Special Issue: Bayesian Inference for Stochastic Processes.

[199] H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, 2009. ISSN 1467-9868. doi: https://doi.org/10.1111/j.1467-9868.2008.00700.x.

[200] A. Ruotoistemaki and V. Mannisto. Improvement of the network level PMS in Finland. In *Fifth International Conference on Managing Pavements*, 2001. ISBN 0971174016.

[201] J. Sadeghi, E. R. Najafabadi, and M. Kaboli. Development of degradation model for urban asphalt pavement. *International Journal of Pavement Engineering*, 18(8):659–667, 2017. doi: https://doi.org/10.1080/10298436.2015.1095912.

[202] E. Sall, G. Erhardt, L. Zorn, D. Tischler, R. Alsup, and N. Nassir. Modeling every hill, bus, traffic signal and car: How SF collaboratively built a citywide dynamic traffic assignment model, 2013. [online] https://www.sfcta.org/sites/default/files/content/IT/SFCHAMP/PDFs/2013_AppCon_DTA_Anyway.pdf. Accessed Jun 24, 2019.

[203] N. Santero, A. Loijos, and J. Ochsendorf. Greenhouse gas emissions reduction opportunities for concrete pavements. *Journal of Industrial Ecology*, 17(6):859–868, 2013. doi: https://doi.org/10.1111/jiec.12053.

[204] N. J. Santero and A. Horvath. Global warming potential of pavements. *Environmental Research Letters*, 4(3):034011, 2009. doi: https://doi.org/10.1088/1748-9326/4/3/034011.

[205] N. J. Santero, E. Masanet, and A. Horvath. Life-cycle assessment of pavements. Part I: Critical review. *Resources, Conservation and Recycling*, 55(9-10):801–809, 2011. doi: https://doi.org/10.1016/j.resconrec.2011.03.010.

[206] N. J. Santero, E. Masanet, and A. Horvath. Life-cycle assessment of pavements Part II: Filling the research gaps. *Resources, Conservation and Recycling*, 55(9-10):810–818, 2011. doi: https://doi.org/10.1016/j.resconrec.2011.03.009.

[207] S. Saxe, G. Casey, P. Guthrie, K. Soga, and H. Cruickshank. Greenhouse gas considerations in rail infrastructure in the UK. *Proceedings of the Institution of Civil Engineers - Engineering Sustainability*, 169(5):171–180, 2016. doi: https://doi.org/10.1680/jensu.15.00015.

[208] S. Saxe, E. Miller, and P. Guthrie. The net greenhouse gas impact of the Sheppard subway line. *Transportation Research Part D: Transport and Environment*, 51:261–275, 2017. doi: https://doi.org/10.1016/j.trd.2017.01.007.

[209] B. Schrödle and L. Held. A primer on disease mapping and ecological regression using INLA. *Computational Statistics*, 26(2):241–258, 2011. doi: https://doi.org/10.1007/s00180-010-0208-2.

[210] G. Scora and M. Barth. Comprehensive Modal Emissions Model (CMEM), version 3.01. User's Guide, 2006. [online] https://www.cert.ucr.edu/cmem/docs/CMEM_User_Guide_v3.01d.pdf. Accessed Feb 2, 2019.

[211] SF Department of Environment. 2017 SF geographic greenhouse gas emisssions inventory at a glance, 2019. [online] https://sfenvironment.org/sites/default/files/fliers/files/sfe_cc_2017_community_inventory_report.pdf. Accessed Jun 24, 2019.

[212] SF Public Works. SF street pavement condition improves for 5th year in a row, 2017. [online] https://sfpublicworks.org/project/press-release-sf-street-pavement-condition-improves-5th-year-row-132017. Accessed Jun 24, 2019.

[213] SFCTA. TNCs Today: A profile of SF transportation network company activity, 2017. [online] https://www.sfcta.org/emerging-mobility/tncs-today. Accessed Jun 24, 2019.

[214] H. Shafer. Why sustainable mobility matters, 2017. [online] https://blogs.worldbank.org/transport/why-sustainable-mobility-matters. Accessed Jun 24, 2019.

[215] M. Y. Shahin and S. D. Kohn. Overview of the PAVERpavement management system and economic analysis of field implementing the PAVER pavement management system. Technical Report ADA116311, Army Construction Engineering Research Laboratory, Champaign, Illinois, USA, 1982. [online] http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA116311. Accessed Jun 24, 2019.

[216] M. Y. Shahin, M. I. Darter, and S. D. Kohn. Development of a Pavement Condition Index for roads and streets. Technical Report ADA057148, Army Construction Engineering Research Laboratory, Champaign, Illinois, USA, 1978. [online] http://www.dtic.mil/dtic/tr/fulltext/u2/a057148.pdf. Accessed Jun 24, 2019.

[217] M. Sheehan. Bay Area UE solver, 2017. [online] http://bit.ly/2RAfOd7. Accessed Jun 24, 2019.

[218] C. Sheppard, R. Waraich, A. Campbell, A. Pozdnukov, and A. R. Gopal. Modeling plug-in electric vehicle charging demand with BEAM: The framework for behavior energy autonomy mobility, 2017. [online] https://escholarship.org/content/qt55p1w1vk/qt55p1w1vk.pdf. Accessed Jun 24, 2019.

[219] B. L. Smith and J. M. Ulmer. Freeway traffic flow rate measurement: Investigation into impact of measurement time interval. *Journal of Transportation Engineering*, 129 (3):223–229, 2003. doi: https://doi.org/10.1061/(ASCE)0733-947X(2003)129:3(223).

[220] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002. doi: https://doi.org/10.1111/1467-9868.00353.

[221] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3):485–493, 2014. ISSN 1467-9868. doi: https://doi.org/10.1111/rssb.12062.

[222] M. Steele. Equilibrium solver, 2012. [online] http://www.repsilat.com/EquilibriumSolver.html. Accessed Jun 24, 2019.

[223] S. Susilawati, M. A. Taylor, and S. V. Somenahalli. Distributions of travel time variability on urban roads. *Journal of Advanced Transportation*, 47(8):720–736, 2013. doi: https://doi.org/10.1002/atr.192.

[224] R. Swan and J. Tucker. Growth projections collide with SF's goal of zero net carbon emissions by 2050, 2019. [online] https://www.sfchronicle.com/bayarea/article/Growth-projections-collide-with-SF-s-goal-of-13868107.php. Accessed Jun 24, 2019.

[225] S. G. Tan and D. Cheng. Quality assurance of performance data for pavement management systems. In *Design, Analysis, and Asphalt Material Characterization for Road and Airfield Pavements*, chapter 5, pages 163–169. American Society of Civil Engineers, 2014. doi: https://doi.org/10.1061/9780784478462.020.

[226] S. Tao, V. Manolopoulos, S. Rodriguez Duenas, and A. Rusu. Real-time urban traffic state estimation with A-GPS mobile phones as probes. *Journal of Transportation Technologies*, 2(1):22–31, 2012. doi: https://doi.org/10.4236/jtts.2012.21003.

[227] C. Tchervenkov, J. Molloy, and K. W. Axhausen. Estimating externalities from GPS traces using MATSim. In *18th Swiss Transport Research Conference (STRC 2018)*. STRC, 2018. [online] https://www.research-collection.ethz.ch/handle/20.500.11850/264806. Accessed Jun 24, 2019.

[228] L. Tierney and J. B. Kadane. Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81(393):82–86, 1986. doi: https://doi.org/10.1080/01621459.1986.10478240.

[229] Transport Research Board. Highway Capacity Manual, 2010.

[230] L. L. Tupper, M. A. Chowdhury, L. Klotz, and R. N. Fries. Measuring sustainability: How traffic incident management through intelligent transportation systems has greater energy and environmental benefits than common construction-phase strategies for "green" roadways. *International journal of sustainable transportation*, 6(5):282–297, 2012. doi: https://doi.org/10.1080/15568318.2011.597910.

[231] UK Road Liaison Group. Surface condition assessment for the national network of roads, 2011. [online] http://www.ukroadsliaisongroup.org/en/asset-condition/road-condition-information/data-collection/scanner/index.cfm. Accessed Jun 24, 2019.

[232] United States Department of Energy. Average fuel economy at different road grades, 2014. [online] https://afdc.energy.gov/data/10601. Accessed Jun 24, 2019.

[233] F. van Wageningen-Kessels, H. Van Lint, K. Vuik, and S. Hoogendoorn. Genealogy of traffic flow models. *EURO Journal on Transportation and Logistics*, 4(4):445–473, 2015. doi: [online]\url{https://doi.org/10.1007/s13676-014-0045-5}.AccessedJun24, 2019.

[234] N. R. Velaga and K. Pangbourne. Achieving genuinely dynamic road user charging: issues with a GNSS-based approach. *Journal of Transport Geography*, 34:243–253, 2014. doi: https://doi.org/10.1016/j.jtrangeo.2013.09.013.

[235] Vital Signs. Street pavement condition, 2017. [online] http://www.vitalsigns.mtc.ca.gov/street-pavement-condition. Accessed Jun 24, 2019.

[236] Vital Signs. Traffic volumes at regional gateways, 2017. [online] http://www.vitalsigns.mtc.ca.gov/traffic-volumes-regional-gateways. Accessed Jun 24, 2019.

[237] M. Wachs. Transportation policy, poverty, and sustainability: history and future. *Transportation Research Record*, 2163(1):5–12, 2010. doi: https://doi.org/10.3141/2163-01.

[238] H. Wang, R. Gangaram, and S. Gresavage. Life cycle assessment of asphalt pavement maintenance, 2014. [online] https://cait.rutgers.edu/wp-content/uploads/2018/05/cait-utc-013-final.pdf. Accessed Jun 24, 2019.

[239] K. C. Wang, J. Zaniewski, and G. Way. Probabilistic behavior of pavements. *Journal of Transportation Engineering*, 120(3):358–375, 1994. doi: https://doi.org/10.1061/(ASCE)0733-947X(1994)120:3(358).

[240] T. Wang, I.-S. Lee, A. Kendall, J. Harvey, E.-B. Lee, and C. Kim. Life cycle energy consumption and GHG emission from pavement rehabilitation with different rolling resistance. *Journal of Cleaner Production*, 33:86–96, 2012. doi: https://doi.org/10.1016/j.jclepro.2012.05.001.

[241] T. Wang, J. Harvey, and A. Kendall. Reducing greenhouse gas emissions through strategic management of highway pavement roughness. *Environmental Research Letters*, 9(3):034007, 2014. doi: https://doi.org/10.1088/1748-9326/9/3/034007.

[242] T. Wang, J. Harvey, J. Lea, and C. Kim. Impact of pavement roughness on vehicle free-flow speed. *Journal of Transportation Engineering*, 140(9):04014039, 2014. doi: https://doi.org/10.1061/(ASCE)TE.1943-5436.0000689.

[243] Y. Wang and N. L. Nihan. Can single-loop detectors do the work of dual-loop detectors? *Journal of Transportation Engineering*, 129(2):169–176, 2003. doi: https://doi.org/10.1061/(ASCE)0733-947X(2003)129:2(169).

[244] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(3):325–362, 1952. doi: https://doi.org/10.1680/ipeds.1952.11259.

[245] World Health Organization. Air polluton, 2019. [online] https://www.who.int/sustainable-development/transport/health-risks/air-pollution/en/. Accessed Jun 24, 2019.

[246] Z. Wu, G. W. Flintsch, and T. Chowdhury. Hybrid multiobjective optimization model for regional pavement-preservation resource allocation. *Transportation research record*, 2084(1):28–37, 2008. doi: https://doi.org/10.3141/2084-04.

[247] X. Xu, J. Gregory, and R. Kirchain. Climate change mitigation potential of pavement albedo, 2018. [online] https://cshub.mit.edu/sites/default/files/documents/ResearchBrief2018_XuAlbedo_FINAL_Linked.pdf. Accessed Jun 24, 2019.

[248] Y. Yin and S. Lawphongpanich. Internalizing emission externality on road networks. *Transportation Research Part D: Transport and Environment*, 11(4):292–301, 2006. [online] https://doi.org/10.1016/j.trd.2006.05.003. Accessed Jun 24, 2019.

[249] B. Yu and Q. Lu. Life cycle assessment of pavement: Methodology and case study. *Transportation Research Part D: Transport and Environment*, 17(5):380–388, 2012. doi: https://doi.org/10.1016/j.trd.2012.03.004.

[250] X. Yu and P. D. Prevedouros. Performance and challenges in utilizing non-intrusive sensors for traffic data collection. *Advances in Remote Sensing*, 2(02):45, 2013. doi: https://doi.org/10.4236/ars.2013.22006.

[251] I. Zaabar and K. Chatti. Calibration of HDM-4 models for estimating the effect of pavement roughness on fuel consumption for US conditions. *Transportation Research Record*, 2155(1):105–116, 2010. doi: https://doi.org/10.3141/2155-12.

[252] W. Zeng, T. Miwa, and T. Morikawa. Prediction of vehicle co2 emission and its application to eco-routing navigation. *Transportation Research Part C: Emerging Technologies*, 68:194–214, 2016. doi: https://doi.org/10.1016/j.trc.2016.04.007.

[253] H. Zhang, G. A. Keoleian, and M. D. Lepech. Network-level pavement asset management system integrated with life-cycle analysis and life-cycle optimization. *Journal of infrastructure Systems*, 19(1):99–107, 2012. doi: https://doi.org/10.1061/(ASCE)IS.1943-555X.0000093.

[254] H. M. Zhang. A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological*, 36(3):275–290, 2002. doi: https://doi.org/10.1016/S0191-2615(00)00050-3.

[255] Y. Zhang and B. Guindon. Using satellite remote sensing to survey transport-related urban sustainability: Part 1: Methodologies for indicator quantification. *International Journal of Applied Earth Observation and Geoinformation*, 8(3):149–164, 2006. doi: https://doi.org/10.1016/j.jag.2005.08.005.

[256] N. Zheng, R. A. Waraich, K. W. Axhausen, and N. Geroliminis. A dynamic cordon pricing scheme combining the macroscopic fundamental diagram and an agent-based traffic model. *Transportation Research Part A: Policy and Practice*, 46(8):1291–1303, 2012. doi: https://doi.org/10.1016/j.tra.2012.05.006.

[257] M. Zhong, R. Shan, D. Du, and C. Lu. A comparative analysis of traditional four-step and activity-based travel demand modeling: a case study of Tampa, Florida. *Transportation Planning and Technology*, 38(5):517–533, 2015. doi: https://doi.org/10.1080/03081060.2015.1039232.

[258] B. Zhou, J. Cao, X. Zeng, and H. Wu. Adaptive traffic light control in wireless sensor network-based intelligent transportation system. In *2010 IEEE 72nd Vehicular Technology Conference-Fall*, pages 1–5. IEEE, 2010. doi: 10.1109/VETECF.2010.5594435.

Appendix A

Determination of the intercity travel demand for San Francisco

As the main source of travel demand data for the case study area does not include trips that start or end outside of San Francisco. Additional information has been collected to estimate this intercity traffic flow between San Francisco and nearby areas.

Geographically, San Francisco pennisula is isolated from the rest of the Bay Area in the north and east directions. Traffic going into and out of the city mainly goes through four "gates": (1) Golden Gate Bridge (to the North); (2) Bay Bridge (to cities on the East); (3) California State Route 1 (SR 1, to the South) and (4) US 101 south and I-280 South (to the San Francisco International Airport, SFO (show in Figure 3.13 in Chapter 3). As a result, the intercity travel demand is determined by collecting the traffic counts information at these four gates. The sources of such information and assumptions involved in data processing are listed here.

1. Golden gate bridge
    - 1,654,022 southbound traffic in November, 2016. Based on http://goldengate.org/news/transit/trends_traffic-transit.php#ggbtrends
    - Assuming southbound and northbound traffic is equal, this is equal to 55,000 daily traffic per direction.
2. Bay bridge
    - In the Fiscal year 2016-17, there are 136,813,538 toll-paid vehicles. Based on https://mtc.ca.gov/about-mtc/what-mtc/mtc-organization/three-agencies-one/bay-area-toll-authority/historic-toll-paid

- Assuming eastbound and westbound traffic is equal, this is equal to 180,000 daily traffic per direction.

3. California State Route 1 (SR1)

   - In 2016, at San Francisco Alemany Boulevard, 103,000 southbound AADT and 96,000 northbound AADT. Based on http://www.dot.ca.gov/trafficops/census/docs/2016_aadt_volumes.pdf
   - Assume northbound and southbound traffic is equal, this is equal to around 100,000 daily traffic per direction.

4. Highway to San Francisco Airport (SFO)

   - In November 2016, SFO total enplaned (departure) is 2,137,473. Total deplaned (arrival) is 2,127,669. Based on http://media.flysfo.com.s3.amazonaws.com/media/sfo/media/air-traffic/as201611.pdf
   - Assume every passenger arrives in individual cars, this is equal to 70k daily traffic per direction.

Appendix B

Detailed results from the sustainability simulations

B.1 10 year average CO_2 emissions

The daily CO_2 emissions over a 10-year analysis period are given in this section for further references. Specifically, the results are given by:

- Scenario 1: PCI-based maintenance without eco-routing.
- Scenario 2: Eco-maintenance without eco-routing.
- Scenario 3: PCI-based maintenance eco-routing. The participation rates of eco-routing are given in the brackets.
- Scenario 4: Eco-maintenance eco-routing. The participation rates of eco-routing are given in the brackets.

Also, the numbers of the local roads and the highways do not add up to the values of whole network, as a small percentage of privately maintained streets are not included in the local roads (maintained by the San Francisco Public Works) and highways (maintained by the Caltrans).

Table B.1 10 year average daily CO2 (t) on the local roads

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	1891	1960	1881	1931
Scenario 2	1879	1926	1869	1894
Scenario 3 (10%)	1883	1951	1874	1924
Scenario 3 (50%)	1857	1919	1851	1899
Scenario 3 (100%)	1870	1922	1867	1913
Scenario 4 (10%)	1873	1919	1863	1889
Scenario 4 (50%)	1849	1896	1844	1877
Scenario 4 (100%)	1866	1912	1866	1908

Table B.2 10 year average daily CO2 (t) on the highways

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	1627	1661	1627	1661
Scenario 2	1627	1661	1627	1661
Scenario 3 (10%)	1605	1641	1605	1640
Scenario 3 (50%)	1526	1568	1522	1560
Scenario 3 (100%)	1436	1490	1430	1470
Scenario 4 (10%)	1606	1641	1604	1639
Scenario 4 (50%)	1526	1565	1521	1553
Scenario 4 (100%)	1433	1478	1423	1450

Table B.3 10 year average daily CO2 (t) of the whole network

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	3820	3934	3810	3905
Scenario 2	3808	3900	3798	3869
Scenario 3 (10%)	3791	3906	3782	3877
Scenario 3 (50%)	3693	3806	3682	3775
Scenario 3 (100%)	3625	3739	3615	3706
Scenario 4 (10%)	3781	3873	3770	3841
Scenario 4 (50%)	3684	3778	3672	3744
Scenario 4 (100%)	3618	3715	3605	3677

B.2 10 year average daily Vehicle Hours Travelled (VHT) and Vehicle Kilometres Travelled (VKMT)

The daily VHT and VKMT over a 10-year analysis period are given in this section for further references. Specifically, the results are given by:

- Scenario 1: PCI-based maintenance without eco-routing.
- Scenario 2: Eco-maintenance without eco-routing.
- Scenario 3: PCI-based maintenance eco-routing. The participation rates of eco-routing are given in the brackets.
- Scenario 4: Eco-maintenance eco-routing. The participation rates of eco-routing are given in the brackets.

Also, the numbers of the local roads and the highways do not add up to the values of whole network, as a small percentage of privately maintained streets are not included in the local roads (maintained by the San Francisco Public Works) and highways (maintained by the Caltrans).

Table B.4 10 year average daily VHT on the local roads

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	337919	337919	337919	337919
Scenario 2	337919	337919	337919	337919
Scenario 3 (10%)	339336	339173	339425	339399
Scenario 3 (50%)	353229	352185	354078	354134
Scenario 3 (100%)	408446	405660	409313	409914
Scenario 4 (10%)	339399	339331	339460	339670
Scenario 4 (50%)	353924	354268	354602	356255
Scenario 4 (100%)	411743	416528	412439	420205

Table B.5 10 year average daily VHT on the highways

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	227642	227642	227642	227642
Scenario 2	227642	227642	227642	227642
Scenario 3 (10%)	223828	224182	223874	223906
Scenario 3 (50%)	213385	214913	212891	213479
Scenario 3 (100%)	212781	216553	211539	213170
Scenario 4 (10%)	223986	224161	223667	223853
Scenario 4 (50%)	213292	214330	212591	212399
Scenario 4 (100%)	212161	214626	210353	209949

Table B.6 10 year average daily VHT of the whole network

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	649074	649074	649074	649074
Scenario 2	649074	649074	649074	649074
Scenario 3 (10%)	648581	648660	648628	648544
Scenario 3 (50%)	680308	681472	681459	681574
Scenario 3 (100%)	814136	814730	810442	814583
Scenario 4 (10%)	648733	648775	648417	648598
Scenario 4 (50%)	681391	682545	681298	681918
Scenario 4 (100%)	817206	826549	812746	820956

Table B.7 10 year average daily VKMT on the local roads

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	6359246	6359246	6359246	6359246
Scenario 2	6359246	6359246	6359246	6359246
Scenario 3 (10%)	6384963	6381521	6387393	6388883
Scenario 3 (50%)	6467365	6442795	6481574	6476545
Scenario 3 (100%)	6552381	6495491	6579328	6569461
Scenario 4 (10%)	6385477	6382638	6389508	6391084
Scenario 4 (50%)	6471089	6456825	6492018	6508287
Scenario 4 (100%)	6563597	6533760	6602552	6639223

B.3 Final PCI of the local roads

The starting condition of the local roads, in terms of the average PCI, is 74. The final PCI of the local roads at the a 10-year analysis period are given in this section for further references. Specifically, the results are given by:

Table B.8 10 year average daily VKMT on the highways

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	8507533	8507533	8507533	8507533
Scenario 2	8507533	8507533	8507533	8507533
Scenario 3 (10%)	8415610	8423760	8413961	8416070
Scenario 3 (50%)	8053928	8093169	8040007	8061169
Scenario 3 (100%)	7450814	7540433	7421217	7464205
Scenario 4 (10%)	8415737	8422664	8411900	8414117
Scenario 4 (50%)	8052751	8079903	8029569	8030285
Scenario 4 (100%)	7437713	7490957	7393598	7380146

Table B.9 10 year average daily VKMT of the whole network

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	15869124	15869124	15869124	15869124
Scenario 2	15869124	15869124	15869124	15869124
Scenario 3 (10%)	15806642	15809970	15806847	15808214
Scenario 3 (50%)	15539471	15547264	15538611	15542365
Scenario 3 (100%)	15017444	15035219	15012200	15022404
Scenario 4 (10%)	15807286	15809158	15806768	15806795
Scenario 4 (50%)	15539537	15546028	15536213	15538098
Scenario 4 (100%)	15012752	15021443	15005969	15003096

- Scenario 1: PCI-based maintenance without eco-routing.
- Scenario 2: Eco-maintenance without eco-routing.
- Scenario 3: PCI-based maintenance eco-routing. The participation rates of eco-routing are given in the brackets.
- Scenario 4: Eco-maintenance eco-routing. The participation rates of eco-routing are given in the brackets.

Table B.10 Final PCI of the local roads at the a 10-year analysis period

Budget, IRI impact	200, 0.01	200, 0.03	700, 0.01	700, 0.03
Scenario 1	73	73	86	86
Scenario 2	70	70	79	79
Scenario 3 (10%)	73	73	86	86
Scenario 3 (50%)	73	73	86	86
Scenario 3 (100%)	73	73	86	86
Scenario 4 (10%)	70	70	79	79
Scenario 4 (50%)	70	70	79	79
Scenario 4 (100%)	70	70	78	77